Three Essays in International Economics

Dissertation
for the Faculty of Economics,
Business Administration and Information Technology
of the University of Zurich

to achieve the title of
Doctor of Philosophy
in Economics

presented by

Andreas Kohler
from Zullwil SO

approved in July 2013 at the request of

Prof. Dr. Josef Zweimüller
Prof. Dr. Fabrizio Zilibotti

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorises the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, July 17, 2013

Chairman of the Doctoral Committee: Prof. Dr. Dieter Pfaff

Contents

Preface

I would like to thank my dissertation committee, Josef Zweimüller and Fabrizio Zilibotti. I especially want to express my deepest gratitude to my thesis supervisor Josef Zweimüller for his support and encouragement. His enthusiasm for economic research is truly inspiring.

My thanks also go to all my colleagues at Professor Zweimüller's chair. In particular, Claudia Bernasconi and Bea Brunner. I couldn't have wished for better company during the last four years. I am further grateful to Andreas Kuhn, Andreas Steinhauer, Christian Hepenstrick, Jean-Philippe Wüllrich, Philippe Ruh, Sandro Favre, Stefan Staubli, Andreas Kettemann, Tobias Renkin, Tobias Würgler, Christian Kiedaisch and Aderonke Osikominu. It wouldn't have been half as much fun without them. In addition, my co-authors Reto Foellmi and Sandra Hanslin deserve special mention for their great collaboration.

I also wish to thank my family and friends. I am indebted to all my friends, especially, Remo, Roger, and Ben. Last, but by no means least, I am grateful to my parents, Marcel and Irene Kohler, for their help and continual support.

Ich danke meinem Referenten Josef Zweimüller und meinem Koreferenten Fabrizio Zilibotti. Mein Dank geht besonders an Josef Zweimüller für seine massgebliche Unterstützung. Seine Begeisterung für die Ökonomie ist inspirierend.

Weiter danke ich dem gesamten Lehrstuhl Zweimüller. Ganz besonders möchte ich Claudia Bernasconi und Bea Brunner für die gemeinsame Zeit am Lehrstuhl danken. Mein Dank gilt ebenfalls Andreas Kuhn, Andreas Steinhauer, Christian Hepenstrick, Jean-Philippe Wüllrich, Philippe Ruh, Sandro Favre, Stefan Staubli, Andreas Kettemann, Tobias Renkin, Tobias Würgler, Christian Kiedaisch und Aderonke Osikominu. Dank geht auch an meine Koautoren Reto Foellmi und Sandra Hanslin für die grossartige Zusammenarbeit.

Ich möchte auch meiner Familie und meinen Freunden danken. Besonderer Dank geht an meine Freunde Remo, Roger und Ben. Schliesslich danke ich meinen Eltern, Marcel und Irene Kohler, für ihre uneingeschränkte Hilfe und Unterstützung.

Andreas Kohler
Windisch
March 2013

1 Introduction

> The price which is so high relatively to the poor man as to be almost prohibitive, may be scarcely felt by the rich; the poor man, for instance, never tastes wine, but the very rich man may drink as much of it as he has a fancy for, without giving himself a thought of its cost (Alfred Marshall, Principles of Economics, 1890, 8th ed., p. 103).

We live in a globalized world marked by the flows of ideas, goods, and capital across national borders. At the same time, we observe large differences in the distribution of income and wealth across and within nations. The distribution of income and wealth across and within countries determines demand patterns, which in turn shape international knowledge, trade and investment flows, and vice versa. This dissertation explores the role of the demand side as a fundamental determinant of these flows.

The *first* essay analyzes the interdependence of innovation incentives and trade patterns in an unequal global economy. It also sets the stage for the second and third essay by providing a theoretical framework that can be adapted to study international product cycles and foreign direct investment. However, first I turn to the question of how differences in per capita incomes across countries shape innovation incentives and trade patterns in the international economy through the demand side. I build upon Grossman and Helpman's (1991b) seminal analysis of innovation and growth in the global economy. I extend their framework by assuming non-homothetic consumer behavior, which allows me to study how the demand structure affects international trade and investment incentives simultaneously. In particular, I analyze a two-country dynamic general-equilibrium model with endogenous growth and costly international trade. In this model arbitrage opportunities in the form of parallel trade emerge if per capita income in one country is sufficiently higher than in the other country. The threat of parallel imports induces some firms located in the country with higher per capita income to forgo a larger market in order to be able to charge higher prices. In other words, households in the relatively poor country cannot afford to purchase all goods produced in the rich country. Thus, less resources in the rich country have to be allocated to the production sector, and more resources can be allocated to research and development. I argue that in an equilibrium where income inequality across countries is high, innovation incentives are relatively high, and at the same time the extensive margin of trade, i.e. the number of different products that are traded, is low. The model allows me to study a variety of interesting issues in the international economy, like trade liberalizations, the design of intellectual property rights, the form of international knowledge spillovers, as well as the effects of an increase in labor productivity in the poor country. As an example, consider the case of a trade liberalization. Trade costs might fall from a level that sustains an equilibrium where all goods are traded to one where only a subset is traded. The reason is that some firms in the rich country start excluding households from the poor country due to the threat of parallel imports. I conclude that a decrease in trade costs might lead to a lower extensive margin of trade but higher incentives to innovate. After the trade liberalization, households in the poor country are unambiguously worse off relative to households in the rich country.

In the *second* essay, joint with Reto Foellmi and Sandra Hanslin, we study the role of the demand side in determining international product cycles. We present a framework that allows us to analyze the effects of an increase in per capita income in emerging countries (e.g. China) relative to developed countries (e.g. United States) on innovation and imitation incentives, as well as trade patterns between those countries. In the 1960s, the American economist Raymond Vernon (1913-1999) proposed the product cycle hypothesis to explain trade patterns between the US and Western Europe (Vernon 1966). His theory offered an explanation that was consistent with trade patterns observed in the data, and which other theories struggled to explain (i.e. Leontief's, 1953, paradox). According to the product cycle theory all products go through the following three stages: In the first stage, a new product is introduced in a country with a (relatively) high per capita income associated with high demand (e.g. the US). As incomes grow abroad, demand for new products emerges in the next advanced countries (e.g. Western Europe), and in the second stage, the product is exported. In the third stage, the product is imitated by firms in less advanced countries, and due to relative cost advantages, production moves there. The more advanced country now imports the product, which it previously exported. Vernon emphasizes the role of the demand channel: "First, the United States market consists of consumers with an average income which is higher . . . than that in any other national market . . . . Wherever there was a chance to offer a new product responsive to wants at high levels of income, this chance would presumably first be apparent to someone in a position to observe the United States market" (Vernon 1966, p. 192). We illustrate the product cycle hypothesis with six major consumer durables, like the microwave or the dishwasher. In particular, we show that across 16 European countries and the US all six consumer durables are introduced earlier in markets with relatively high average incomes. Furthermore, in 1978 the US starts out as a net exporter of all six consumer durables and ends up as a net importer in 2006. For the microwave and the washing machine we observe that US production declines relative to the production in emerging countries like Brazil, Russia, India and China between 1982 and 2008. Motivated by those findings, we formalize the product cycle theory in a dynamic general-equilibrium model. A wealthy North demands and innovates new products, while a poor South at random imitates products made in the North. This contrasts to the first essay where both regions where at the technological frontier, developing new goods. By introducing non-homothetic preferences the model is able to generate endogenous product cycles as hypothesized by Vernon (1966), and observed in the data. Besides technology, the different stages of the product cycle are now determined by the distribution of income between

North and South. We argue that a reduction in income inequality across North and South leads to a decline in the innovation and imitation rate. Since Northern households become poorer the incentives to innovate in the North diminish. However, since households in the South are richer, firms in the North want to export their products sooner so that the first stage of the product cycle becomes shorter. At the same time, the average duration a new product is manufactured in the North increases because imitation activity in the South is less intense. This implies that the third stage of the cycle, when products are exported from South to North, becomes shorter. We further discuss extensions with respect to preferences and technology.

In the first and second essay, firms have no choice but to export if they want to serve a foreign market. However, besides international trade a globalized economy is characterized by investment flows across national borders. The *third* essay studies the choice of firms between foreign direct investment (FDI) and exports in general equilibrium. The market size hypothesis, as formulated e.g. by Balassa (1966), argues that FDI will take place if the market is large enough to capture economies of scale. This essay analyzes the market size hypothesis from a demand-side perspective. I argue that in the presence of a proximity-concentration trade-off, i.e. to take advantage of economies of scale firms want to concentrate production while at the same time locate production close to their consumers to avoid transportation costs, a firm's foreign market entry mode depends on the market segment it serves. In particular, firms serving mass markets abroad engage in FDI whereas firms selling exclusively to a few rich consumers abroad export. This implies that the distribution of income within a country is important in determining the country's attractiveness for FDI. I formalize these arguments in a simple static general-equilibrium model with two regions Home and Foreign, and low, middle and high income classes of consumers in both regions. I think of Home as the wealthy region relative to Foreign. Consumers have non-homothetic preferences, and need to satisfy subsistence consumption in terms of food before they spend all additional income on differentiated consumer goods. In equilibrium, ex-ante identical firms choose different pricing strategies depending on the market segment (i.e. income classes) they cater to. Firms producing differentiated products for the mass market (i.e. for middle and rich classes in both regions) engage in FDI whereas firms serving exclusively the rich classes in both regions export. The basic intuition is that for firms supplying the mass market economies of scale are high enough to compensate the higher fixed costs associated with FDI. Due to the assumptions about the distribution of income across and within regions, products sold on the mass market are priced according to the willingness to pay of the middle class in Foreign, which is determined by their average income. This is the highest price a firm can set if it wants to sell on the mass market. Hence, I ask how changes in the average income of Foreign's middle class affect the number of producers engaged in FDI, ceteris paribus. I argue that redistributing income within Foreign towards its middle class increases the equilibrium number of multinational producers with headquarters in Home. Furthermore, I show that an expansion in the size of the middle class in Foreign, ceteris paribus, has ambiguous effects on FDI activity in Foreign, depending on whether the poor or rich class shrinks. The model is extended along various dimension like technology, preferences, and differentiated factor endowments. As an illustration, I use data on foreign direct investment of OECD countries. I construct common income classes for a sizeable number of countries using inequality data, and apply different definitions of a global middle class used in the literature. In line with the baseline model I find a positive relationship between the average income of the middle class in a host country and its FDI position held by OECD countries.

2 Innovation and Trade in an Unequal Global Economy

2.1 Introduction

We observe vast differences in per capita incomes across countries and regions in the world. For example in 2009, GDP per capita in the United States was about six times as high as GDP per capita in China. In 2009, even average GDP per capita across all OECD countries was about three times as high as average GDP per capita across all non-OECD countries (PWT 7.0, PPP, 2005 USD). Instead of analyzing the sources of theses differences in per capita incomes across countries, this chapter looks at the other side of the same coin, by providing a theoretical framework that allows us to study the consequences of differences in per capita incomes across countries or regions for the patterns of international trade and the incentives to innovate in the world economy.

To this end, we extend Grossman and Helpman's (1991b) seminal analysis of trade and growth in a global economy by addressing the effects of income inequality across countries on international trade and innovation. Our contribution to the literature is to introduce non-homothetic preferences in Grossman and Helpman's (1991b) general-equilibrium model of horizontal endogenous innovation and costly international trade between two countries. This allows us to study the channels through which the structure of the demand side affects the incentives to innovate in an international economy, and at the same time, determines the patterns of trade between countries. In the standard model with homothetic preferences inequality in per capita incomes across countries has no effect on the incentives to innovate or the patterns of international trade. The reason is that with homothetic preferences only total lifetime income in the world economy matters for innovation and trade but not its distribution across countries. However, with non-homothetic preferences the distribution of income across countries has important implications for the incentives to innovate and the patterns of international trade. We show that in the presence of transportation costs, arbitrage opportunities emerge if the level of inequality across countries is relatively high. This threat of parallel imports induces some firms located in the rich country, i.e the country with the relatively high per capita income, to forgo a larger market (market size effect) in order to be able to charge higher prices (price effect). Since households in the poor country cannot afford to consume all goods produced in the rich country less resources have to be allocated to the production sector to satisfy the needs of households in the poor country. Hence, more resources can be allocated to the research and development sector of the economy, which allows the (world) economy to grow at a higher rate. An equivalent argument is that when incomes are concentrated in one country, firms can charge high markups, and therefore the incentives to innovate are high. In contrast to the standard model, where markups are determined by the exogenously given elasticity of substitution between goods, markups are endogenously determined by the willingness to pay of households, and hence by the distribution of income. Furthermore, unlike in the standard model firms cannot pass on trade costs so that markups on products sold abroad are lower than on products sold at home. In sum, at high levels of inequality not all goods produced are traded internationally, which is reminiscent of the Burenstam-Linder (1961) hypothesis, while at the same time, incentives to innovate are high.

The introduction of non-homothetic preferences naturally raises a number of interesting issues. For example, we can use our framework to discuss the design of intellectual property rights across countries. In particular, we argue that households in the poor country might not see eye to eye with households in the rich country on whether there should be national or international exhaustion of patents. We show that under a policy of national exhaustion compared to one of international exhaustion households in the rich country incur only losses since their consumption grows at a lower rate relative to a policy of international exhaustion, whereas households in the poor country incur an additional static gain since they can consume a larger share of products. Whether the households in the two countries agree or disagree about patent policy depends crucially on how poor households weigh current gains against future losses.

Furthermore, we can analyze the consequences of a trade liberalization on trade patterns, the incentives to innovate, and relative welfare levels. We contend that a trade liberalization might increase the incentives to innovate, decrease trade at the extensive margin, and make households in the poor country worse off relative to households in the rich country. If trade costs fall from a level that sustains an equilibrium where all goods are traded to one that sustains an equilibrium where only a subset of all goods are traded, some firms in the rich country start to exclude households from the poor country due to the threat of parallel imports. This implies that resources are released from the production sector, which have previously been used to satisfy the consumption needs of households in the poor country, that can now be allocated to the research sector. Hence, the world economy grows at a higher rate but at the same time the extensive margin of trade falls. One can then show that poor households become unambiguously worse off relative to rich households.

Grossman and Helpman (1991b) initiated a vast literature that studies the incentives to innovate in a global economy in general-equilibrium models with homothetic preferences. The literature that looks at the relationship between income inequality and growth, and in particular at the channels through which income inequality affects economic growth is much smaller. In particular, it focuses on credit market imperfections (e.g. Galor and Zeira 1993) or political economy mechanisms (e.g. Persson and Tabellini 1994; Alesina and Rodrik 1994) but largely neglects demand-side effects. There are relatively few models that do consider demand-side effects, and they can be broadly divided into two categories. On the one hand, there are static models of international trade and non-homothetic preferences as in Foellmi et al. (2011), Markusen (2010), Matsuyama (2000) or Mitra and Trindade (2005) that focus on the effects of the demand side on trade patterns. On the other hand, there are dynamic models that analyze how the demand side shapes the incentives to innovate in closed economies like Foellmi and Zweimüller (2006).

This model combines the static open economy setup from Foellmi et al. (2011) with the dynamic structure of Foellmi and Zweimüller (2006). In particular, the model nests the two models, i.e. it collapses to Foellmi and Zweimüller (2006) with no consumption hierarchy in the case of no trade costs and integrated factor markets, and to Foellmi et al. (2011) in the case where the time horizon becomes arbitrarily small. This allows us to create a framework where the distribution of income across countries affects trade patterns and the incentives to innovate simultaneously. Hence, we can study the interdependence of trade and innovation in a global economy and compare our results to Grossman and Helpman (1991b) where the distribution of income across countries is irrelevant due to the assumption of homothetic preferences.

We proceed in our analysis as follows. We start in Section 2.2 by examining the closed economy with a focus on the difference between non-homothetic and standard CES preferences. Section 2.3 turns to the open economy and studies in detail the different equilibria as well as the conditions under which they emerge. We analyze the effects of changes in income inequality across countries on the incentives to innovate, and the patterns of international trade in Section 2.4. To illustrate the transitional dynamics, we look at the example of a labor productivity surge in research and development in the poor country, which could be the result of a market or education reform. Section 2.5 compares the model to Grossman and Helpman (1991b) in some detail. We then turn to applications and extensions in Section 2.6, and analyze how the design of intellectual property rights, trade policy, and the form of international knowledge spillovers affects trade patterns and the innovation incentives. In the case of trade policy we also study the welfare consequences of a trade liberalization. Last, we analyze how general our results are with respect to the assumptions about preferences in the baseline model. Section 2.7 concludes.

2.2 Closed Economy

In this section we first discuss in detail the closed economy equilibrium with a representative household to gain intuition for the impact of preferences on innovation before turning to the open economy case.

2.2.1 Distribution and Endowments

There is a total of L identical households in the economy, which trivially allows for a representative household. Suppose that each household inelastically supplies θ efficiency units of labor in the labor market, and holds assets $a(t)$ (i.e. shares in firm's profits). Hence, per capita income in period t is given by $y(t) = w(t)\theta + r(t)a(t)$, where $w(t)$ denotes the wage rate per efficiency unit, and $r(t)$ the return on assets (i.e. interest rate).

2.2.2 Households

Preferences

The household's utility function is defined over a continuum of differentiated and indivisible goods $j \in [0, \infty)$. At any point in time only a finite subset $N(t)$ is available in the market. All households have the same non-homothetic instantaneous utility function given by

$$u\left(\{c(j,t)\}_{j=0}^{\infty}\right) = \int_{j=0}^{\infty} v\left(c\left(j,t\right)\right) dj \tag{2.1}$$

where baseline utility $v\left(c\left(j,t\right)\right)$ is an indicator function $c\left(j,t\right) \in \{0,1\}$, and we normalize $v(0) = 0$ and $v(1) = 1$. This utility function satisfies the standard properties of positive and non-increasing marginal utility since utility from consuming the first unit is one but utility from consuming a second unit is zero, i.e. households are (locally) satiated. The marginal utility from consuming an infinitesimal (in this case, zero) amount of any good j is finite, i.e. $\lim_{c(j,t)\to 0} u'\left(\cdot\right) = 1 < \infty$. Hence, non-negativity constraints $c\left(j,t\right) \geq 0$ might become binding for some j's. Consequently, wealthy households will consume a different consumption bundle than poor households. For example, suppose that at some point in time the available set of indivisible goods consists of the following durables, a car, a washing machine, and a television set. In that case, affluent households can afford to purchase all durables whereas poor households might only afford to buy a television. Put differently, households can only choose how many different goods they consume (extensive margin) but not how much of each good they consume (intensive margin). In this sense, zero-one preferences are the antipode of, and no less general than, constant-elasticity-of-substitution (CES) preferences, where households can only choose consumption along the intensive margin. In Section 2.6.5 we discuss quadratic preferences where households have a choice along the intensive and extensive margin of consumption. Last, note that all goods enter the utility function (2.1) symmetrically, i.e. there is no consumption hierarchy. A consumption hierarchy could be incorporated in the utility function by introducing a weighting function $\xi(j)$, which is decreasing in the index j (see Foellmi and Zweimüller, 2006). This would imply that the marginal utility from consuming good j is higher than from consuming good k, where $j < k$. Hence, households would first purchase low-indexed goods and as their incomes grow move on to higher-indexed goods. This would make the analysis substantially more complex without adding much insight to how inequality across countries affects trade and growth. However, it would allow us to make a statement about which country specializes in high-, respectively, low-indexed goods.

Household Problem

We assume that a household has a logarithmic intertemporal utility function given by

$$U(0) = \int_{t=0}^{\infty} \exp(-\rho t) \log u\left(\{c(j,t)\}_{j=0}^{\infty}\right) dt \tag{2.2}$$

where $u(\cdot)$ is given by (2.1), and $\rho > 0$ denotes the time preference rate.[1] The household maximizes (2.2) subject to non-negativity constraints $c(j,t) \geq 0$, $\forall j,t$, and its intertemporal budget constraint

$$\int_{t=0}^{\infty} \exp\left(R(t)\right) e(t)dt \leq a(0) + \int_{t=0}^{\infty} \exp\left(R(t)\right) w(t)\theta dt \tag{2.3}$$

where $p(j,t)$ denotes the price of good j at time t, $a(0)$ initial wealth of the household, $R(t) = \int_{s=0}^{t} r(s)ds$ the cumulative interest rate, and $e(t) = \int_0^\infty p(j,t)c(j,t)dj$ consumption expenditures. We imposed a no-Ponzi game condition of the following form $\lim_{t\to\infty} \exp\left(R(t)\right) a(t) \geq 0$ on the intertemporal budget constraint. The first-order conditions (including complementary slackness conditions) to the household's optimization problem are given by

$$\begin{aligned}
e^{-\rho t}\frac{u'(\cdot)}{u(\cdot)} - \Lambda \exp\left(R(t)\right) p(j,t) + \mu(j,t) &= 0 && (2.4)\\
\mu(j,t)c(j,t) &= 0,\ \mu(j,t) \geq 0,\ c(j,t) \geq 0 && (2.5)\\
\int_{t=0}^{\infty} \exp\left(R(t)\right) e(t)dt &= a(0) + \int_{t=0}^{\infty} \exp\left(R(t)\right) w(t)\theta dt && (2.6)\\
\lim_{t\to\infty} \exp\left(R(t)\right) \lambda(t)a(t) &= 0 && (2.7)
\end{aligned}$$

where Λ denotes the (present value) Lagrange multiplier on the intertemporal budget constraint, and $\mu(j,t)$ the Lagrange multiplier on the non-negativity constraints. Note that the current value Lagrange multiplier on the budget constraint is given by $\lambda(t) = \Lambda \exp\left(-R(t) + \rho t\right)$. Due to the transversality condition the intertemporal budget constraint will always be binding in optimum, i.e. preferences exhibit global non-satiation. We distinguish the following cases:

(i) Non-negativity constraint is binding, i.e. $c(j,t) = 0$. This implies by (2.5) that $\mu(j,t) \geq 0$. Hence, the first-order condition (2.4) can be written as follows

$$\exp(-\rho t)\frac{u'(\cdot)}{u(\cdot)} \leq \Lambda \exp\left(-R(t)\right) p(j,t) \tag{2.8}$$

where the left-hand side denotes the marginal utility gain from consuming good j and the right-hand side denotes the marginal utility cost from consuming good j.

(ii) Non-negativity constraint is not binding, i.e. $c(j,t) > 0$. This implies by (2.5) that $\mu(j,t) = 0$. Therefore, the first-order condition (2.4) can be written as follows

$$\exp(-\rho t)\frac{u'(\cdot)}{u(\cdot)} = \Lambda \exp\left(-R(t)\right) p(j,t). \tag{2.9}$$

[1]Notice that logarithmic preferences are the special case of constant-relative-risk-aversion (CRRA) preferences of the following form $u(\cdot)^{1-\sigma}/(1-\sigma)$, where σ denoting the inverse of the intertemporal elasticity of substitution goes to 1. Since the focus of our analysis lies on the consequences of inequality across countries on intra-temporal instead of inter-temporal consumption decisions we choose analytical convenience over generality and let $\sigma \to 1$.

From equations (2.8) and (2.9) we can derive the household's Marshallian demand function for good j as

$$c\,(j,t) = \begin{cases} 0, & p(j,t) > z(j,t) \\ 1, & p(j,t) \leq z(j,t) \end{cases} \tag{2.10}$$

where $z(j,t) \equiv [u(\cdot)\Lambda \exp\left(-R(t)+\rho t\right)]^{-1}$ denotes the household's willingness to pay for some good j. Intuitively, households compare the marginal utility gain of consuming 1 unit of good j, i.e. $\exp(-\rho t)u'(\cdot)/u(\cdot)$, with its associated marginal utility cost, i.e. $\Lambda \exp\left(-R(t)\right) p(j,t)$. If the marginal utility gain exceeds or is equal to the marginal utility cost households purchase 1 unit of good j, otherwise they don't. In equilibrium $u(\cdot)^{-1} = N(t)^{-1}$ will hold, where $N(t)$ denotes the number (measure) of differentiated goods of which the household consumes one unit each. We see that the willingness to pay depends negatively on the Lagrange multiplier Λ, which can be interpreted as the marginal utility of lifetime wealth. The wealthier a household, the lower its marginal utility of wealth Λ, and therefore, holding everything else constant, the higher its willingness to pay $z(j,t)$. In other words, if a household becomes wealthier its willingness to pay for a given good will increase if it cannot expand its consumption to new goods, i.e. increase $N(t)$. This is the case because households have no choice about how much they want to consume of each good (intensive margin). Furthermore, the more different goods a household consumes, i.e. the higher $N(t)$, the lower its (marginal) willingness to pay for a given good j, ceteris paribus.

Summing over individual household demands (2.10) yields aggregate or market demand for good j

$$C(p(j,t)) = \begin{cases} 0, & p(j,t) > z(j,t) \\ L, & p(j,t) \leq z(j,t). \end{cases} \tag{2.11}$$

The step function in Figure 2.1 below depicts aggregate demand for good j. From Figure 2.1 we can derive the price elasticity of aggregate demand $\varepsilon \equiv -C'(p(j,t))p(j,t)/C(p(j,t))$ as follows. On the one hand, if firm j increases its price $p(j,t)$ to a price above $z(j,t)$ its demand drops to zero, implying that aggregate demand is infinitely price elastic for $p(j,t) > z(j,t)$, i.e. the price elasticity $\varepsilon \to \infty$. On the other hand, if firm j increases its price to any price below or equal to $z(j,t)$ its demand remains unchanged. Hence, aggregate demand is perfectly inelastic for $p(j,t) < z(j,t)$, i.e. the price elasticity $\varepsilon \to 0$. In other words, for price changes below $z(j,t)$ households always consume 1 unit of good j regardless of its price, which is an implication of local satiation in households' preferences.

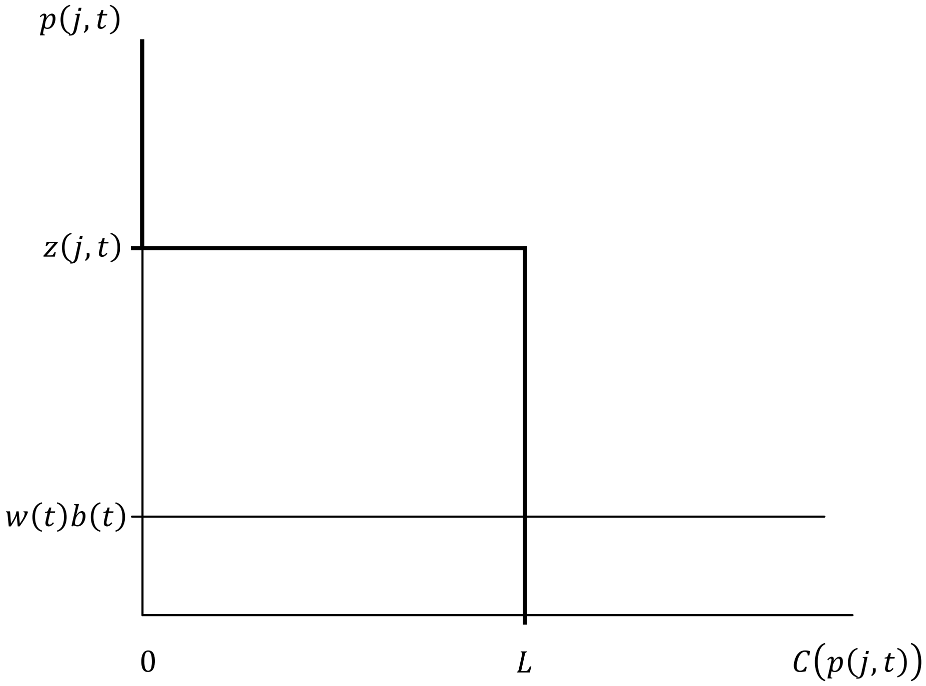


Figure 2.1: Aggregate demand for good j in period t

2.2.3 Firms

Technology

Firms have access to the following technology. A firm must hire $F(t) = F/N(t)$ units of labor in the labor market up-front to do research and development (R&D) for a new product, where $F > 0$ is a positive constant. Each firm will develop and produce only one unique good. Firms raise the necessary capital to finance their R&D projects by issuing shares in the capital market. Remember that $N(t)$ denotes the stock of products, or knowledge imbedded in these products, which have been developed in the economy up to time t. This implies that there are intertemporal knowledge spillovers. As in Romer (1990) ideas are non-excludable and non-rival. In other words, the more products have been developed in the past the more productive is labor in the research lab ("standing on the shoulders of giants"). Once a new product has been developed the production of 1 unit of output requires $b(t) = b/N(t)$ units of labor to produce, with $b > 0$. We assume that intertemporal spillovers carry over to the production sector. Finally, new products are protected by patents which grant the innovating firms perpetual monopolies.

Firm Problem

Firm j maximizes operating profits

$$\pi(j,t) = [p(j,t) - w(t)b(t)]\, C(p(j,t)) \tag{2.12}$$

subject to its aggregate demand $C(p(j,t))$ given by (2.11). The first-order condition is given by

$$\frac{p(j,t) - w(t)b(t)}{p(j,t)} = \frac{1}{\varepsilon} \tag{2.13}$$

where the left-hand side denotes the Lerner index, and the right-hand side denotes the inverse of the price elasticity of aggregate demand. It follows that the markup, i.e. the ratio of price $p(j,t)$ to marginal cost $w(t)b(t)$, is determined by $\mathcal{M} \equiv \left(1 - \frac{1}{\varepsilon}\right)^{-1}$. We make two observations. First, the monopolist always operates in a price region where the price elasticity of aggregate demand ε exceeds 1. In the region where the elasticity is less than 1, the monopolist's revenue and his profits are increasing in price and decreasing in quantity (see e.g. Tirole 1988). Second, the markup $\mathcal{M}$ is decreasing in the price elasticity of aggregate demand ε.

From the first-order condition (2.13) of firm j we can derive the optimal price as follows. Suppose firm j sets a price $p(j,t)$ larger than $z(j,t)$. In that case, aggregate demand for good j becomes zero, i.e. the price elasticity of aggregate demand ε goes to infinity. Thus, the markup $\mathcal{M}$ converges to 1, and firm j does not earn any (economic) profits needed to recoup the fixed cost $w(t)F(t)$, paid up-front for product development. Hence, firm j would never increase its price above the willingness to pay $z(j,t)$ and run losses. Consider the case where firm j sets a price $p(j,t)$ below $z(j,t)$. In that price region the monopolist does not operate because aggregate demand is perfectly inelastic, i.e. the price elasticity ε is zero. Thus, the highest price firm j can set is equal to the willingness to pay $z(j,t)$. This is the price at which marginal revenue equals marginal cost. We conclude that the profit-maximizing price for firm j is given by

$$p(j,t) = z(j,t)\,. \tag{2.14}$$

With CES preferences the markup $\mathcal{M}$ is exogenously determined by the elasticity of substitution between goods whereas in the case of zero-one preferences the markup is endogenously given by $\mathcal{M} = z(j,t)/w(t)b(t)$. Given marginal cost, the markup increases if the willingness to pay of households increases. This implies that the higher the willingness to pay, the lower the price elasticity of demand, ceteris paribus.

2.2.4 Equilibrium

Goods Market

In equilibrium, a finite set $N(t)$ of differentiated goods is available on the market. From optimal prices (2.14) follow optimal quantities in equilibrium. Since all firms face the same demand curve (2.1) and have the same cost structure, they all charge the same price $p(t) = z(t)$ equal to the willingness to pay of households, and supply the same quantity given by

$$C(t) = L,\ \forall t. \tag{2.15}$$

From the first-order condition (2.9) follows that optimal consumption expenditures grow at the following rate

$$\frac{\dot{e}(t)}{e(t)} = r(t) - \rho \tag{2.16}$$

where $u'(\cdot)/u(\cdot) = 1/N(t)$ since $c(t) = 1$ for all j's consumed in equilibrium. The Euler equation says that nominal consumption expenditures of the household $e(t) = z(t)N(t)c(t)$, where $c(t) = 1$, grows at a positive rate if the marginal utility gain from postponing consumption $r(t)$ outweighs its marginal utility cost ρ. Notice that the intensive margin of consumption is constant over time, i.e. $c(t) = 1$ for all t.

Labor Market

The labor market clears in every period. Because labor is mobile across sectors, i.e. all households can work in the production or the R&D sector, there is one wage rate at which the labor market clears. In other words, the following arbitrage argument applies. If the wage rate in the production sector is higher than in the R&D sector all workers flock to the production sector, and vice versa. This rises the marginal product of labor in the R&D sector. Hence, wages in the R&D sector increase such that workers move back. The movement of workers across sectors goes on until wages are equalized across sectors. Hence, the wage rate is determined by relative labor demand from the production to the R&D sector. Labor market clearing is determined by

$$\begin{aligned} L\theta &= \dot{N}(t)F(t) + \int_{j=0}^{N(t)} b(t)X(j,t)dj = g(t)F + bL \\ L\theta &= g(t)F + bL \end{aligned} \tag{2.17}$$

where $g(t) = \dot{N}(t)/N(t)$. The left-hand side of (2.17) denotes exogenous aggregate labor supply, and the right-hand side aggregate labor demand from the R&D and production sector. Labor demand from the production sector is determined only by technology parameter b and population size L due to the constant intensive margin of consumption, i.e. $c(j,t) = 1$ for all j, t. In particular, this implies that equilibrium labor demand from the production sector is independent of prices $p(t)$, and hence the willingness to pay $z(t)$.

We see that the economy faces the following fundamental trade-off between consumption today and tomorrow. Higher consumption today requires a larger share of labor allocated to the production sector in the economy. However, this implies that a lower share of labor can be allocated to the research sector and therefore less new products are developed. Hence, consumption tomorrow is lower.

Capital Market

We assume perfect capital markets. On perfect capital markets, the present discounted value of profits equals the stock market value of the firm. Since all firms make identical profits the present discounted value of profits is determined by

$$v(t) = \int_{\tau=t}^{\infty} \exp\left(-R(\tau)\right) \pi(\tau) d\tau \tag{2.18}$$

where equilibrium profits in period t are determined by $\pi(t) = [z(t) - w(t)b(t)] L$. Differentiating (2.18) with respect to time t yields the familiar arbitrage condition

$$\pi(t) + \dot{v}(t) = r(t)v(t) \tag{2.19}$$

which says that the return on an investment of size $v(t)$ in any firm, i.e. the dividend $\pi(t)$ paid plus capital gains/losses $\dot{v}(t)$, must equal the return on an investment of size $v(t)$ in risk-less bonds. We assume that there is free entry into product development. There is firm-entry as long as the value of a firm $v(t)$ is equal to or exceeds the fixed cost of product development $w(t)F(t)$. Hence, in an equilibrium with positive growth free entry implies the following zero-profit condition

$$v(t) = w(t)F(t). \tag{2.20}$$

Capital market clearing requires that savings equal investment, i.e. aggregate asset holdings $A(t) = La\,(t)$ must equal the stock market value of firms $\int_{j=0}^{N(t)} v\,(t)\, dj = N(t)v(t)$. The interest rate $r(t)$ adjusts such that the capital market clears in every period t.

2.2.5 Steady State

We consider a steady state in which all variables grow at the same constant rate $g > 0$. Hence, prices, quantities, profits, the interest rate, and the shares of labor allocated to production and R&D are constant in the steady state.

Equations (2.15)-(2.20) completely characterize the steady state. We choose the marginal cost of production as the numeraire $w(t)b(t) = 1$. The equations can be reduced to the labor market clearing condition and the zero-profit condition in the endogenous variables g and z:

$$L\theta = gF + bL \tag{2.21}$$

$$\frac{(z-1)L}{g+\rho} = \frac{F}{b}. \tag{2.22}$$

Remember that because the intensive margin of consumption is constant the share of labor allocated to the production sector is determined entirely by technology and population size in equilibrium, and the residual share of labor is allocated to R&D. Hence, the solution for the growth rate of the economy g is pinned down by the labor market clearing condition

$$g = \frac{1}{F/bL}\left(\frac{\theta}{b} - 1\right) > 0 \;\Leftrightarrow\; \theta > b. \tag{2.23}$$

In other words, the growth rate g is determined such that labor demand and labor supply equalize, i.e. the labor market clears. It follows that equilibrium prices $z = (L\theta + F\rho)/bL$ and markups adjust such that the zero-profit condition holds, i.e. are consistent with free entry into product development.

2.2.6 Discussion

In this discussion we focus on the consequences of endogenous markups in our model. This highlights the difference to the standard model of Grossman and Helpman (1991b) with CES preferences where markups are exogenously determined.

We start by observing that the growth rate is independent of the preference parameter ρ. In the steady state, an increase in ρ has too opposing effects on the growth rate g, which cancel each other. On the one hand, an increase in ρ leads to an increase in the interest rate r. This tends to decrease the growth rate g through its negative effect on the present discounted value of profits. On the other hand, an increase in ρ leads to an increase in markups and prices z, which tends to increase profits and therefore the growth rate g. In the steady state, these two effects offset each other. To understand the intuition behind this result consider a ceteris paribus increase in the time preference rate ρ, i.e. households become more impatient. In other words, households prefer even more to consume today instead of tomorrow. As in standard models of endogenous growth, if households are less willing to save the interest rate r needed to equate savings with investments must increase. In the model with CES preferences where markups are exogenously determined by the elasticity of substitution between goods, a change in ρ leaves markups and prices unchanged. The increase in r then leads to a lower present discounted value of profits, and hence a lower incentive to innovate. However, in this model, markups are endogenously determined by the willingness to pay of households. If households become more impatient their willingness to pay for goods today relative to tomorrow increases such that markups and prices rise. This effect exactly offsets the effect through the interest rate.

2.3 Open Economy

In this section, we turn to the case of the open economy. Suppose that there are two countries or regions indexed by $i = \{R, P\}$. We allow goods to be traded across countries at iceberg trade costs $\tau \geq 1$. This implies that to ensure one unit of a good arrives at its destination, $\tau \geq 1$ units of that good have to be shipped.

2.3.1 Distribution and Endowments

There is a total of L households in the two countries. A fraction β of the total population lives in country P whereas the rest $1 - \beta$ lives in country R. Each household in country i inelastically supplies θ_i efficiency units of labor in their domestic labor market. In particular, we assume that labor is immobile across countries but is mobile across sectors within a country. We further assume that each household holds only domestic assets, i.e. there is a perfect home bias in their portfolios. Hence, inequality in the endogenously determined personal income distribution *across* countries originates from differences in labor and capital incomes across countries. We will assume that households in country R are wealthier than households in country P. Therefore, we sometimes refer to country R as the rich, and country P as the poor country.

There exists an institution which levies a lump-sum tax $T_R(t) > 0$ in the rich country in each period and redistributes the tax revenues as lump-sum benefits $T_P(t) > 0$ to households in the poor country. We impose that the institution runs a balanced budget in each period, i.e. total tax revenues $(1-\beta)LT_R(t)$ equal total tax spending $\beta LT_P(t)$. Hence, we can choose the level of $T_P(t) = T(t)$ as our exogenous variable, which implies that $T_R(t) = \beta T(t)/(1-\beta)$. We assume that over time transfers grow at the same constant rate g as incomes, i.e. $T(t) = T(0)e^{gt}$. This assumption is necessary to generate a steady state in which the distribution of income across countries is stationary. Changes in lump-sum transfers will allow us to study ceteris paribus changes in inequality across countries, i.e. without changing relative country sizes, labor endowments or technology at the same time. We measure income inequality across countries in the steady state according to the Gini coefficient.

2.3.2 Households

All households maximize intertemporal utility (2.2), where instantaneous utility $u(\cdot)$ is given by (2.1), subject to their intertemporal budget constraint (2.3) and non-negativity constraints $c_i(j,t) \geq 0$. All households have the same preferences regardless of their country of residence. The resulting market demand for good j is given by

$$X(p(j,t)) = \begin{cases} 0, & p(j,t) > z_R(j,t) \\ (1-\beta)L, & z_P(j,t) < p(j,t) \leq z_R(j,t) \\ L, & p(j,t) \leq z_P(j,t) \end{cases} \tag{2.24}$$

where $z_i(j,t)$ denotes the willingness to pay of households in country i for good j. Note that because households in country R are wealthier than households in country P their marginal utility of wealth $\lambda_R(t)$ is lower. This implies that they have a higher willingness to pay for good j, i.e. $z_R(j,t) > z_P(j,t)$, holding the set of goods consumed $N_i(t)$ constant. Furthermore, wealthy households have a relatively low price elasticity of demand. Figure 2.2 shows market demand (2.24) for good j.

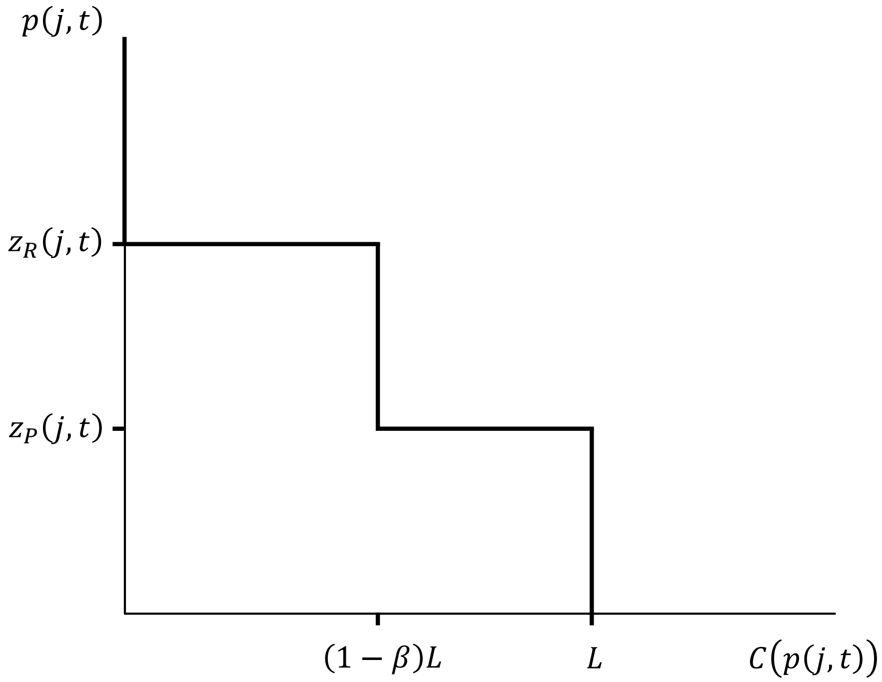


Figure 2.2: Aggregate demand for good j in period t

2.3.3 Firms

Technology

Firms located in country i have access to the following technology, which we allow to differ across countries. A firm must hire $F^i(t) = F^i/N(t)$ units of labor in the domestic labor market up-front to do research and development (R&D) for a new product. Each firm will develop and produce only one unique good. Firms raise the necessary capital to finance their R&D projects by issuing shares in their domestic capital market. Note that $N(t)$ denotes the stock of products which have been developed in the world economy up to time t, i.e. $N(t) = \sum_i N^i(t)$. This implies that there are international and intertemporal knowledge spillovers like in Grossman and Helpman (1991b). Coe and Helpman (1995) find empirical evidence for the existence of international R&D spillovers. Their estimates suggest that foreign R&D has beneficial effects on domestic productivity, and that these effects are stronger the more open an economy is to foreign trade. In Section 2.6.4 we relax this assumption, and look at the consequences of imperfect international knowledge spillovers.[2] Once a new product has been developed the

[2] As in Grossman and Helpman (1991b), trade, i.e. the integration of product markets, introduces competition among innovators across countries. This implies that firms in both countries have an incentive to develop new products that are unique in the world. Therefore, trade eliminates duplication of research and development across countries. To see this, suppose that a firm in country i, with access to the technology described above, develops a product that already exists in country $k \neq i$. Ex-post price competition between the entrant and the incumbent would result in a limit price equal to the highest marginal cost, at that point the firm with the relatively higher marginal cost drops out of the market. Hence, the monopoly profits a firm earns from capturing an existing market, i.e. copying an existing product, are strictly lower than the monopoly profits from opening up a new market, i.e. developing a novel product. In other words, the strategy of imitating an existing product is always strictly dominated by the strategy to develop a novel product.

production of 1 unit of output requires $b^i(t) = b^i/N(t)$ units of domestic labor to produce. New products are protected by internationally exhausted patents which grant the developing firms a perpetual monopoly. Notice that internationally exhausted patents imply that parallel imports are legal. In Section 2.6.1 we take a closer look at the design of international property rights, and their effects on the incentives to innovate and trade patterns. Furthermore, we abstract from foreign direct investment or licensing of technology to foreign firms.

Firm Problem

Firm j maximizes profits (2.12) subject to its aggregate demand (2.24) and subject to a price setting restriction implied by the international exhaustion of patents, i.e. the threat of parallel imports. The profit-maximizing behavior of firm j is stated in Proposition 2.1 below.

Proposition 2.1. *In the case, where (i) the willingness to pay of households in the rich country, $z_R(j,t)$ falls short of the willingness to pay of households in the poor country $z_P(j,t)$ times trade costs τ, firms can perfectly price discriminate across countries. In the case, where (ii) the willingness to pay in the rich country $z_R(j,t)$ exceeds or is equal to the one in the poor country $z_P(j,t)$ times trade costs τ, the price-setting power of firms is limited.*

Proof. Let us first prove case (ii) of the proposition. Suppose that $z_R(j,t) \geq \tau z_P(j,t)$ with $\tau > 1$. Firm j located in country R would like to choose prices in the rich and poor country such that its profits

$$\pi_R(j,t) = \left[p_R^R(j,t) - w^R(t)b^R(t)\right](1-\beta)L + \left[p_P^R(j,t) - \tau w^R(t)b^R(t)\right]\beta L \tag{2.25}$$

are maximized, where $w^R(t)$ denotes the wage rate in the rich country, $p_R^R(j,t)$ the price of good j in the rich country, and $p_P^R(j,t)$ the price in the poor country. The first term in (2.25) denotes profits earned on the domestic market and the second term profits earned from serving the foreign market. We assume $p_P^R(j,t) \geq \tau w^R(t)b^R(t)$. Otherwise, firm j in the rich country would make a loss on each unit it exports to the poor country and would therefore never export in the first place. Recall our discussion in Section 2.2.3 about optimal price setting of monopolistic firms. In the export market firm j would like to set a price equal to the willingness to pay of households in the poor country, and in the domestic market a price equal to the willingness to pay of households in the rich country. Suppose, firm j would set a price in the domestic market equal to $z_R(j,t)$ and in the export market equal to $z_P(j,t)$. Since $z_R(j,t) \geq \tau z_P(j,t)$ arbitrageurs could make positive profits by buying good j on the grey market in the poor country at low prices $z_P(j,t)$, and importing it back to the rich country at cost τ, where they could sell it at a price marginally lower than $z_R(j,t)$. This threat forces firm j to set a limit price equal to $\tau z_P(j,t)$ in the rich country if it doesn't want to loose its domestic demand to arbitrageurs. The same argument holds for firm j producing in the poor country and exporting to the rich country. Hence, if $z_R(j,t) \geq \tau z_P(j,t)$ the firm's price setting power is limited by the threat of parallel imports, i.e. the price setting restriction imposed by internationally exhausted patents becomes binding. The proof of part (i) is straightforward.

If $z_R(j,t) < \tau z_P(j,t)$, the arbitrage described above is not profitable, and therefore firms can perfectly price discriminate across countries. Consider the special case of no trade cost, i.e. $\tau = 1$. Since $z_R(j,t) > z_P(j,t)$ at any point in time, firm j engaged in international sales is always restricted in its price setting power due to the threat of parallel imports. Hence, firm j exporting its good is forced to set a limit price of $z_P(j,t)$ in the rich country, and case (i) never occurs. □

2.3.4 Equilibrium

Based on Proposition 2.1 we make the following proposition, which determines the structure of the equilibrium.

Proposition 2.2. *Proposition 2.1 implies that in case (ii) not all firms producing in the rich country export to the poor country, and in case (i) all firms export in equilibrium.*

Proof. Suppose that in case (ii) all firms would sell in the rich country at prices $\tau z_P(t)$. Since $\tau z_P(t) < z_R(t)$ for all goods j this implies that households in the rich country would not exhaust their budget constraints. Hence, their willingness to pay for an additional good would be infinitely high. This would induce some firms producing in the rich country to sell their goods exclusively in their domestic market. Since these firms don't export they are not subject to the threat from parallel imports, and can set prices equal to the willingness to pay of rich households. In equilibrium, there exist two types of firms producing in the rich country, the ones that export, and the ones that don't, where both types of firms must earn the same profits. Since firms manufacturing in the poor country have to incur transportation costs if they export, it is never a dominant strategy for these firms to sell exclusively in the rich country. In case (i), the price setting restriction does not bind and no firm has an incentive to sell exclusively in the rich country, given exporting is profitable in the first place. □

Let us refer to case (i) as a *full*-trade equilibrium, and case (ii) as a *part*-trade equilibrium. In other words, if income inequality across countries is high for a given level of trade costs, or trade costs are low for a given level of inequality, only part of all goods produced in the rich country will be traded, and the economy is in a part-trade equilibrium. Note that the equilibrium structure is identical to the static model in Foellmi et al. (2011). For a comparison of the equilibrium structure to the standard model see Section 2.5 below.

Parallel Imports

The model suggests that parallel imports and price discrimination across countries on the basis of differences in per capita incomes are important characteristics in the open economy. First, both casual observation and anecdotal evidence suggest that parallel imports might be important in reality. Maskus and Chen (2004) state that it is difficult to measure how important parallel imports are since they are generally not recorded. However, according to them, survey evidence suggests that they can capture a sizable share of markets for specific products where parallel trade is allowed. Furthermore, there are numerous case studies which suggest that parallel imports are far from irrelevant in reality. For example, the consultancy KPMG (2003, 2008), estimated the size of the IT grey market at about eight per cent of total global IT sales. In the Economist (1998), among other examples, the car maker Honda claims to have lost as much as a quarter of its British sales to parallel imports in 1998, and the British supermarket chain Tesco has been enjoined from selling Levi's jeans cheaper than authorized sellers. Furthermore, Simonovska (2010) provides empirical evidence that per capita incomes and prices of homogeneous tradable goods are positively related.

2.3.5 Full-trade Equilibrium

In this section, we discuss case (i) implied by Proposition 2.2. In a full-trade equilibrium, all households in both countries consume all goods available on the world market. However, households in the rich country pay higher prices for the same consumption bundle. In other words, firms fully skim the higher willingness to pay of rich households.

Goods Markets

In equilibrium, a finite set $N(t) = N^P(t)+N^R(t)$ of differentiated goods is available on the world market, where $N^i(t)$ denotes the set of differentiated goods produced in country i. Proposition 2.2 implies that in a full-trade equilibrium all goods sold in country i command the same prices, equal to the willingness to pay $z_i(t)$ of households in country i, regardless of where they are produced. Hence, all firms supply the following quantities in equilibrium

$$C(t) = L, \ \forall t. \tag{2.26}$$

From the first-order condition (2.9) follows the Euler equation for households in country i

$$\frac{\dot{e}_i(t)}{e_i(t)} = r^i(t) - \rho \tag{2.27}$$

where $e_i(t) = N_i(t)z_i(t)$ denotes nominal consumption expenditures, $N_i(t)$ the set of goods consumed, and $r^i(t)$ the interest rate in country i. In equilibrium, we have $N_i(t) = N(t)$ for all households in both countries.

Labor Markets

Remember that labor is immobile across countries but mobile across production and R&D sectors within countries. Labor market clearing in the rich country is determined by

$$(1-\beta) L\theta_R = g^R(t) [1-m(t)] F^R + [1-m(t)] b^R (1-\beta) L + [1-m(t)] \tau b^R \beta L \tag{2.28}$$

where $g^R(t) \equiv \dot{N}^R(t)/N^R(t)$, and $m(t) \equiv N^P(t)/N(t)$ such that $1-m(t) = N^R(t)/N(t)$ denotes the production share of the poor and rich country, respectively. The left-hand side denotes aggregate labor supply in the rich country, whereas on the right-hand side, the first term denotes labor demand from the R&D sector, and the second and third term labor demand from the production sector supplying domestic and export markets. Similarly, in the poor country, labor market clearing is given by

$$\beta L\theta_P = g^P(t)m(t)F^P + m(t)b^P\beta L + m(t)\tau b^P (1-\beta) L \tag{2.29}$$

where $g^P(t) \equiv \dot{N}^P(t)/N^P(t)$.

Capital Markets

We assume that capital is immobile across countries. Capital markets in each country are perfect such that the present discounted value of profits $v_i(t)$ equals the stock market value of domestic firms. Since all firms within country i have the same demand and cost functions they all make identical profits $\pi_i(t)$. Free entry into product development in country i implies the following zero-profit condition

$$v_i(t) = w^i(t)F^i(t). \tag{2.30}$$

Capital market clearing in country i requires that domestic savings equal domestic investment, i.e. aggregate asset holdings $A_i(t)$ must equal the stock market value of firms $N^i(t)v_i(t)$. The interest rate $r^i(t)$ adjusts such that the capital market in country i clears in every period.

Balance of Payments

We assume that the balance of payments is balanced period by period. For details see Appendix 2.A.1. Since we assume that capital is immobile across countries, i.e. foreign ownership of domestic assets is not allowed, the capital account, which keeps track of the net change in national ownership of assets, is always balanced. Hence, the current account must also be balanced in every period.[3] In other words, we require that the sum of the balance of trade plus net transfer payments is zero in every period. Hence, the balance of payments is given by

$$\left[\beta LN^R(t)z_P(t) - (1-\beta)LN^P(t)z_R(t)\right] - \beta LT_P(t) = 0 \tag{2.31}$$

where the first term in brackets on the left-hand side denotes the balance of trade and the second term net transfer payments. If there is a progressive transfer $T_P(t) > 0$ it must be that the balance of trade is negative for the poor country, i.e. the value of its imports exceeds the value of its exports. The poor country then runs a permanent trade deficit financed by transfers from the rich country.

Steady State

We consider a steady state in which all variables in both countries grow at the common constant rate $g > 0$. In each country prices, quantities, profits, the interest rate, and the share of labor

[3] The current account is the sum of the balance of trade, net factor payments and net transfer payments. Since capital and labor are not mobile across countries net factor payments are zero.

allocated to production and R&D are constant in the steady state. Interest rates equalize across countries since incomes grow at the same rate in each country. Equations (2.26)-(2.31) completely determine the full-trade steady state. We choose the marginal cost of production for firms located in the poor country as the numeraire and set $w^P(t)b^P(t) = 1$ for all t. For completeness all equations describing the steady state are stated in Appendix 2.A.1.

The steady state growth rate g and the production share of the poor country m are determined by the intersection of the labor market clearing conditions (2.28) and (2.29). We label the labor market condition in the rich country *RR-curve*:

$$(1-\beta)\, L\theta_R \;=\; g\,(1-m)\, F^R + (1-m)\, b^R\,(1-\beta)\, L + (1-m)\,\tau b^R \beta L.$$

The labor market condition in the poor country is labelled *RP-curve*:

$$\beta L\theta_P \;=\; gmF^P + mb^P\beta L + m\tau b^P\,(1-\beta)\, L.$$

Figure 2.3 depicts the graphical solution of the steady state. To guarantee that a unique equilibrium exists we make the following assumption:

Assumption 2.1. $-\frac{\beta\left(\theta_P/b^P\right)}{\beta+\tau(1-\beta)} < \frac{(1-\beta)\left(\theta_R/b^R-1\right)-\tau\beta}{(1-\beta)+\tau\beta}$.

Proposition 2.3. *Given Assumption 2.1 holds, a unique steady state with a positive growth rate* $g > 0$ *and production share of the poor country* $0 < m < 1$ *exists.*

Proof. The *RP-curve* is a decreasing convex function, and the *RR-curve* is an increasing concave function in the (m, g)-space. If Assumption 2.1 holds, the y-axis intercepts are such that $RP(m)|_{g=0} > RR(m)|_{g=0}$. □

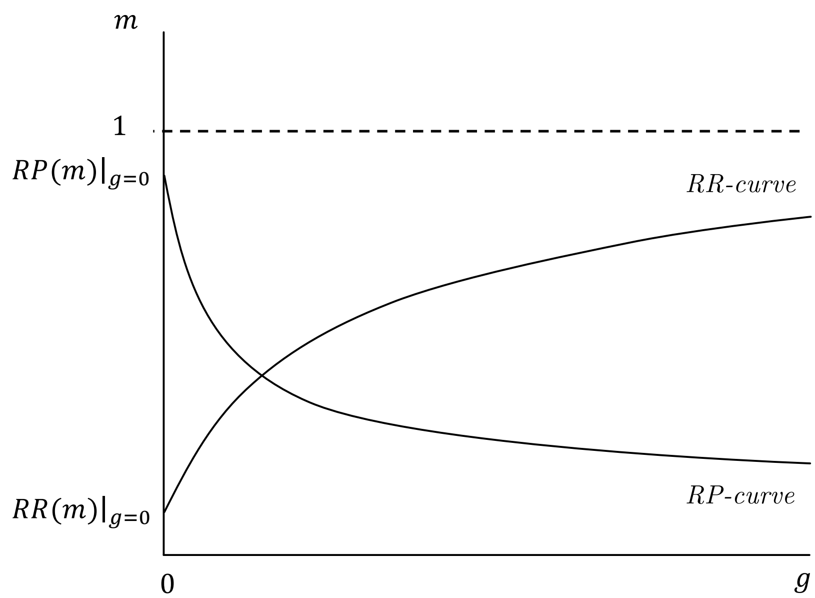


Figure 2.3: Steady state full-trade equilibrium

2.3.6 Part-trade Equilibrium

This section discusses case (*ii*) implied by Proposition 2.2. Households in the rich country consume all goods available on the world market whereas households in the poor country consume only a subset of all goods available. Firms cannot perfectly price discriminate across countries due to the imminent threat of parallel imports. Some firms located in the rich country don't export, i.e. they forgo a larger market in order to be able to charge higher prices (unconstrained price-setting), whereas all firms located in the poor country export (constrained price-setting). Note that we keep all the assumptions made in the full-trade equilibrium concerning the economic environment.

Goods Markets

In equilibrium, a finite set $N(t) = N^R(t) + N^P(t)$ of differentiated goods is available on the market in the rich country whereas only a subset $N_P(t) < N(t)$ is available on the market in the poor country.

Proposition 2.2 implies that in a part-trade equilibrium goods are priced as follows. All goods sold in the poor country command the same prices equal to the willingness to pay of households in the poor country $z_P(t)$. In the rich country, all goods that are not exported are priced at the willingness to pay of households in the rich country $z_R(t)$, whereas all goods that are sold in the rich and poor country, regardless of where they are produced, command prices equal to trade costs times the willingness to pay of households in the poor country $\tau z_P(t)$. Hence, all firms located in the rich country that don't export supply the following quantities in equilibrium

$$C(t) = (1 - \beta)L, \ \forall t. \tag{2.32}$$

All firms that engage in international trade, regardless of their location, supply the following quantities in equilibrium

$$C(t) = L \ \forall t. \tag{2.33}$$

From the first-order condition (2.9) follows the evolution of optimal consumption expenditures $e_i(t)$ for households in country i

$$\frac{\dot{e}_i(t)}{e_i(t)} = r^i(t) - \rho. \tag{2.34}$$

Labor Markets

Labor market clearing in the part-trade equilibrium is determined as follows. In the rich country, labor market clearing is given by

$$(1 - \beta)L\theta_R = g^R(t)\left[1 - m(t)\right]F^R + \left[1 - m(t)\right]b^R(1 - \beta)L + \left[n(t) - m(t)\right]\tau b^R \beta L \tag{2.35}$$

where $n(t) \equiv N_P(t)/N(t)$ denotes the consumption of a household in the poor relative to the rich country. The share of exporting firms in the rich country is defined by $q(t) \equiv N_M^R(t)/N^R(t)$, and can be written as $q(t) = [n(t) - m(t)]/[1 - m(t)]$. We can interpret $q(t)$ as the extensive margin of trade. Labor market clearing in the poor country is determined by

$$\beta L\theta_P = g^P(t)m(t)F^P + m(t)b^P\beta L + m(t)\tau b^P(1-\beta)L \tag{2.36}$$

which is identical to the full-trade equilibrium.

Capital Markets

Proposition 2.2 implies that in a part-trade equilibrium firms located in the rich country must be indifferent whether to export or not. In other words, profits of firms that export $\pi_{R,M}(t)$ must equal profits of firms that don't export $\pi_{R,E}(t)$. Due to free entry into R&D in country i the present discounted value of profits $v_i(t)$ must equal R&D cost $w^i(t)F^i(t)$. Capital market clearing in country i requires that domestic savings equal domestic investment.

Balance of Payments

We require the balance of payments to balance period by period

$$\left[\beta LN_M^R(t)z_P(t) - (1-\beta)LN^P(t)\tau z_P(t)\right] - \beta LT_P = 0 \tag{2.37}$$

where the term in brackets denotes the balance of trade and the second term net transfer payments.

Steady State

We consider a steady state in which all variables in both countries grow at the same constant rate $g > 0$. This implies that in each country prices, quantities, profits, the interest rate, and the share of labor allocated to production and R&D are constant, as well as the share of exporters in the rich country. Equations (2.32)-(2.37) characterize the steady state. We keep the marginal cost of production for firms in the poor country as the numeraire. All equations describing the steady state are stated in Appendix 2.A.2.

The steady state growth rate g and the production share of the poor country m are determined by the intersection of the labor market clearing conditions (2.35) and (2.36). The *RR-curve* is given by:

$$\begin{aligned}(1-\beta)L\theta_R &= g\,(1-m)\,F^R + (1-m)\,b^R(1-\beta)L + m\tau^2 b^R(1-\beta)L \\ &+ \frac{\tau b^R\beta LT\,[\beta+\tau(1-\beta)]}{(g+\rho)\,(F^P/b^PL) + [\beta+\tau(1-\beta)]}\end{aligned} \tag{2.38}$$

where we used the balance of payment condition and the zero-profit condition in the poor country to solve for the share of exporting firms in the rich country q as a function of m and

g. The *RP-curve* is identical to the full-trade equilibrium, and determined by:

$$\beta L\theta_P = gmF^P + mb^P\beta L + m\tau b^P(1-\beta)L. \tag{2.39}$$

Assumption 2.2. *If the following assumptions hold, the RR-curve has a negative slope in the (m,g)-space:*

$$\left(F^P\rho/b^PL\right) + [\beta + \tau(1-\beta)] > (1-\beta)\left(\tau^2-1\right), \textit{and } \ 1 < \theta_R/b^R < \tau^2.$$

If the following assumption holds, the intercept of the RR-curve with the y-axis lies above the intercept of the RP-curve with the y-axis in the (m,g)-space:

$$\frac{\beta\left(\theta_P/b^P\right)}{[\beta+\tau(1-\beta)]} < \frac{\left(\theta_R/b^R-1\right)}{(\tau^2-1)} + \frac{\tau b^R\beta LT\left[\beta+\tau(1-\beta)\right]}{(F^P\rho/b^PL)+[\beta+\tau(1-\beta)]}.$$

Proposition 2.4. *Given Assumption 2.2 holds, a unique steady state with a positive growth rate* $0 < g < \xi$*, where* $\xi = \left(\tau^2-1\right)b^R(1-\beta)L/F^R$*, and production share of the poor country* $0 < m < 1$ *exists.*

Proof. If Assumption 2.2 holds, the RP-curve is a decreasing, convex function, and the RR-curve is a decreasing, concave function in the (m,g)-space, with y-axis intercepts $RR(m)|_{g=0} > RP(m)|_{g=0}$. □

The graphical solution of the steady state is shown in Figure 2.4. The analytical study of the steady state is algebraically complex, so that in the following we will study the properties of the full- and part-trade steady state by simulating the model and looking at small changes in parameters.

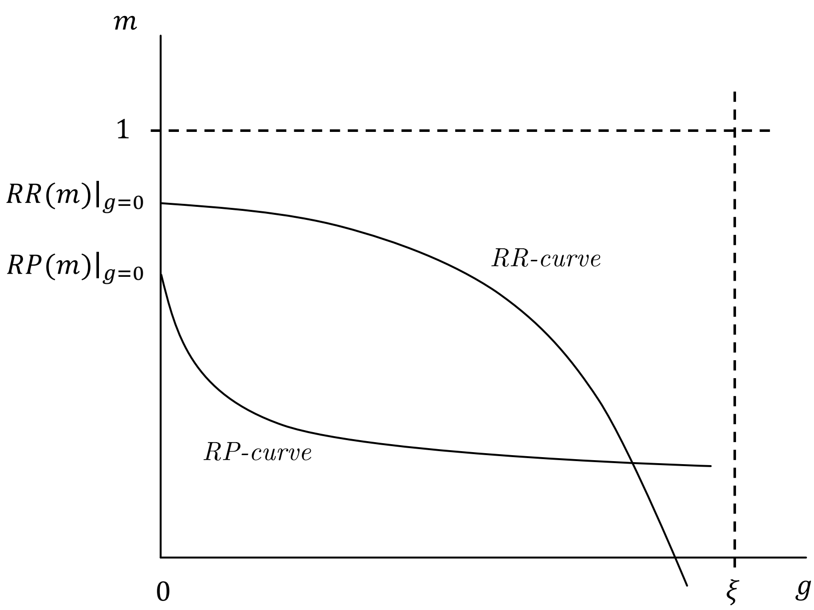


Figure 2.4: Steady state part-trade equilibrium

2.4 Comparative Statics and Transitional Dynamics

In this section we discuss how changes in the distribution of income across countries, and a market or education reform in the poor country that makes labor in the domestic research sector more productive affects the trade patterns and the incentives to innovate. We use the example of a market reform to discuss how the global economy adjusts to a new steady state after a change in the poor country's labor productivity in the research sector occurred.

2.4.1 Change in Income Inequality across Countries

This section discusses the consequences of a change in global, i.e. between-country, inequality for the incentives to innovate and the patterns of international trade. We start our analysis by simulating a full-trade steady state and make regressive transfers, i.e. transfers from the poor country to the rich country, until the economy is in a part-trade steady state. In Appendix 2.A.3 we specify the parameter values used in our simulation and show the simulation results in Figures A.1-A.4. The results are summarized below. Note, that in this section we compare steady states. We will turn to the transitional dynamics in the next section.

Our findings can be summarized as follows: An increase in inequality through a regressive transfer has a positive effect on the growth rate g, and the terms of trade for the poor country, a negative effect on the production share of the poor country m, the share of exporting firms in the rich country q, and the consumption share of households in the poor country n, and has no effect on the relative wage rate.

The intuition is the following. As long as the economy is still in a full-trade steady state after the regressive transfer, i.e. $z_R < \tau z_P$ holds, households in the poor country can still afford to buy one unit of all goods available on the world market. Hence, the incentives to innovate have not changed. In other words, the resource constraints are not affected since there is no intensive margin of consumption. Relative prices z_R/z_P, i.e. the terms of trade, increase but as long as $z_R/z_P < \tau$ holds no firm in the rich country has an incentive not to export. However, if the economy switches to a part-trade steady state after the regressive transfer, i.e. $z_R \geq \tau z_P$ holds, households in the poor country are too poor to afford all goods available on the world market, and relative consumption n falls, ceteris paribus. Some firms in the rich country have an incentive not to export to the poor country so that q decreases. Hence, in the production sector of the rich country resources are set free that can be reallocated to R&D in the rich country. Innovation incentives in the rich country increase. This implies that more new products are being developed in the rich country than in the poor country such that the production share of the poor country m declines. This absorbs additional resources in the production sector of the rich country, which tends to decrease growth, while at the same time, releases resources from the production sector to the R&D sector in the poor country, which tends to increase growth. In our simulations, the positive effects on the growth rate g dominate. Alternatively, one could argue as follows. An increase in inequality across countries, ceteris paribus, increases the willingness to pay of households in the rich relative to the poor country. Hence, the price elasticity of demand of households in the rich country decreases so that firms can charge higher markups on marginal costs and earn higher profits from selling to them. At the same time, households in the poor country become more price sensitive so that markups and profits earned from selling to them decrease. As long as the inequality does not increase too much, and the economy remains in a full-trade steady state, these two effects offset each other. Thus, profits and therefore the incentives to innovate are not affected. However, if inequality increases sufficiently, the economy ends up in a part-trade steady state where households in the poor country can no longer afford to buy all products available (n decreases). In that case, firms that choose not to export can charge higher markups if inequality increases, and therefore earn higher profits. Arbitrage in the rich country implies that profits for firms that do export must increase as well so that firms pursuing either strategy earn the same profits in equilibrium. Remember that the marginal willingness to pay of households in the poor country increases, ceteris paribus, as they consume less differentiated products. Hence, the incentives to innovate increase as inequality across countries increases. As inequality increases, less and less firms in the rich country have an incentive to export (q falls), and due to a home market effect makes relatively more firms locate in the rich country (m decreases). Note that markups on domestically sold goods are higher than on exported goods since trade costs cannot be passed on to consumers.

As inequality increases the terms of trade z_R/z_P move in favor of the poor country until the economy is in a part-trade steady state when the terms of trade are pinned down by trade costs $\tau > 1$ because the price setting restriction $z_R > \tau z_P$ is binding. The intuition is straightforward. As households in the rich country become relatively richer they can pay higher prices z_R whereas households in the poor country become poorer and can pay lower prices z_P. Hence, firms located in the poor country get relatively higher prices for their export goods. Remember that as long as the economy is in a full-trade steady state households in the rich and poor country always buy the same consumption basket but households in the rich country pay higher prices for that same basket as firms can fully skim their willingness to pay. Hence, a regressive transfer increases (decreases) the willingness to pay of households in the rich (poor) country and therefore increases (decreases) prices, holding the number of differentiated goods purchased constant.

Since technologies and population sizes do not change because of a regressive transfer and are still identical across countries, relative wages are always equalized regardless of whether the economy is in a full- or part-trade steady state (see Appendix 2.A.1 and 2.A.2)

In sum, an increase in (lifetime) income inequality across countries leads to an increase in the world growth rate g, and at the same time, to a decrease in the number of differentiated goods that are traded between countries. In other words, the extensive margin of trade depends negatively on inequality across countries. Among others, Bernasconi (2013) finds evidence for the extensive (and intensive) margin of trade, i.e. the higher the overlap of the income distributions of two countries, the more product categories (extensive margin) they trade with each other. We are not aware of an empirical study that looks at the effect of cross-country income inequality on (global) economic growth. The empirical studies we are aware of analyze the relationship between income inequality within countries and growth. The results from these studies are inconclusive. Some, like Barro (2000), find a negative relationship between income inequality within countries and growth for poor countries and a positive relationship for rich countries. Others, like Perotti (1996), find a negative relationship.

2.4.2 Change in Technology

In this section, we discuss the effects of a decline in the labor requirement to create a new product in the poor country F^P. One could think of this case as a market reform, i.e. easier access to product and/or financial markets, or an education reform in the poor country that makes labor in the research and development sector relatively more productive. The results of the numerical simulation are summarized below. The parameter values used in the simulation and Figures A.4-A.6 that show the simulated effects of a decline in F^P are given in Appendix 2.A.4.

We find the following: An increase in labor productivity in R&D in the poor country, i.e. a decrease in F^P, leads to an endogenous decrease in inequality across countries. At high levels of F^P, such that the economy is in a part-trade equilibrium, a reduction of F^P leads to a decrease in the growth rate g, whereas at low levels of F^P, such that the economy is in a full-trade equilibrium, a reduction of F^P leads to an increase in g. The production share of the poor country increases, whereas the relative wage rate and the terms of trade decrease as F^P falls.

Two effects are driving these results. First, if labor is more productive in R&D in the poor country the cost of innovation, ceteris paribus, falls so that the incentives to innovate increase. Second, as labor in R&D becomes more productive in the poor country inequality across countries endogenously decreases. This has a negative effect on the incentives to innovate because households in the poor country can afford to buy more products, which absorbs more resources in the production sector of the rich country, ceteris paribus. In a part-trade equilibrium the second effect outweighs the first effect so that an increase in the labor productivity in the R&D sector of the poor country has a negative effect on the growth rate g. As long as the economy is in a full-trade equilibrium, only the first effect is present. Because households in the poor country consume all goods available inequality has no effect on the incentives to innovate. Hence, the increase in labor productivity $1/F^P$ has a positive effect on the growth rate g.

Since the innovation cost decreases as F^P falls, relatively more new products are developed in the poor country so that m increases, ceteris paribus. As F^P falls and labor in the poor country becomes relatively more productive the relative wage rate $w^R(t)/w^P(t)$ falls. Furthermore, as households in the rich country become relatively less rich, i.e. inequality across countries decreases, the terms of trade fall.

In sum, we get a "U-shape" relationship between labor productivity in R&D of the poor country, which we can think of as a measure of market development in the poor country, and the growth rate of the global economy.

2.4.3 Transitional Dynamics

We illustrate transitional dynamics with the example of a change in technology in the poor country discussed in the previous section. To this end, we discuss the transition for the case of no transfer system, i.e. $T(t) = 0$ for all t. This is convenient since it allows us to solve analytically and graphically for the transitional dynamics. In the full-trade equilibrium transitional dynamics are governed by equations (2.26)-(2.31) and in the part-trade equilibrium by equations (2.32)-(2.37). The labor market clearing conditions (2.28)-(2.29) for the full-trade equilibrium and (2.35)-(2.36) together with the balance of payments (2.37) for the part-trade equilibrium describe a system of autonomous and homogeneous linear first-order differential equations we can write as follows

$$\dot{N}^R(t) = \Phi_k^R N^R(t) + \Psi_k^R N^P(t) \tag{2.40}$$
$$\dot{N}^P(t) = \Phi_k^P N^P(t) + \Psi_k^P N^R(t) \tag{2.41}$$

for $k = \{full, part\}$. The definition of the constants Φ_k^R, Ψ_k^R, Φ_k^P and Ψ_k^P, and the analytical solution to this system are given in Appendix 2.A.5. Here, we analyze the transitional dynamics in the phase diagram. In the steady state, the set of differentiated goods in both countries grows at a constant common rate $g > 0$. For notational convenience we drop subscript k. From equation (2.40) follows $\mu \equiv \left(g - \Phi^R\right)/\Psi^R > 0$ where $\mu \equiv N^P(t)/N^R(t) > 0$ in steady state. Note that parameter values must be such that $\mu(t) > 0$ for all t. Hence, we see that

$$\frac{\dot{N}^R(t)}{N^R(t)} \begin{cases} > g, & \text{if} \quad \mu(t) > \mu \\ = g, & \text{if} \quad \mu(t) = \mu \\ < g, & \text{if} \quad \mu(t) < \mu. \end{cases}$$

Similarly, from equation (2.41) follows $\mu \equiv \Psi^P/\left(g - \Phi^P\right) > 0$. Therefore, we observe that

$$\frac{\dot{N}^P(t)}{N^P(t)} \begin{cases} < g, & \text{if} \quad \mu(t) > \mu \\ = g, & \text{if} \quad \mu(t) = \mu \\ > g, & \text{if} \quad \mu(t) < \mu. \end{cases}$$

Obviously, the steady state growth rate g is determined by $\left(g - \Phi^R\right)\left(g - \Phi^P\right) = \Psi^R\Psi^P$. In sum, if $\mu(t) < \mu$, the set of differentiated goods in the rich country $N^R(t)$ grows at a lower rate than the steady state growth rate g while the set of differentiated goods in the poor country $N^P(t)$ grows at a higher rate than g. Hence, the ratio $\mu(t) = N^P(t)/N^R(t)$ converges monotonically to its steady state value μ, and vice versa, if $\mu(t) > \mu$. The balance of payments condition in the full-trade steady state defines a critical level of $\widetilde{\mu} \equiv \beta/(1-\beta)\tau > 0$, where $\widetilde{\mu}$ satisfies $z_R/z_P = \beta/(1-\beta)\widetilde{\mu} = \tau$. Hence, for $\mu(t) > \widetilde{\mu}$ the dynamics of the economy are governed by the dynamics in a full-trade equilibrium, and vice versa.

Now, we can illustrate the transitional dynamics with the help of the example discussed in Section 2.4.2. In Figure 2.5 we analyze the effect of an exogenous decrease in the labor requirement to develop new products in the poor country F^P, i.e. labor in the R&D sector becomes more productive, which shocks the economy out of its part-trade steady state. We assume that the new level of F^P sustains a full-trade steady state with a higher ratio of firms located in the poor to the rich country $\mu(t)$. During the transition from the part- to the full-trade steady state the ratio μ_{part} monotonically increases until it reaches μ_{full}. As long as $\mu(t) \leq \widetilde{\mu} < \mu_{full}$ the economy is in a part-trade equilibrium, and as $\widetilde{\mu} < \mu(t) \leq \mu_{full}$ the economy is in a full-trade equilibrium. Hence, during the transition households in the poor country must invest relatively more in the development of new products than households in the rich country, i.e. aggregate consumption expenditures of households in the poor country $E_P(t) = \beta Le_P(t)$ must grow at a lower rate during the transition than consumption expenditures of households in the rich country $E_R(t) = (1-\beta)Le_R(t)$. In other words, relatively more resources are temporarily invested in R&D in the poor relative to the rich country. Regardless of whether the economy ends up in full- or part-trade steady state with a higher or lower (world) growth rate, a market reform in the poor country triggers temporarily more investment in R&D in the poor country.

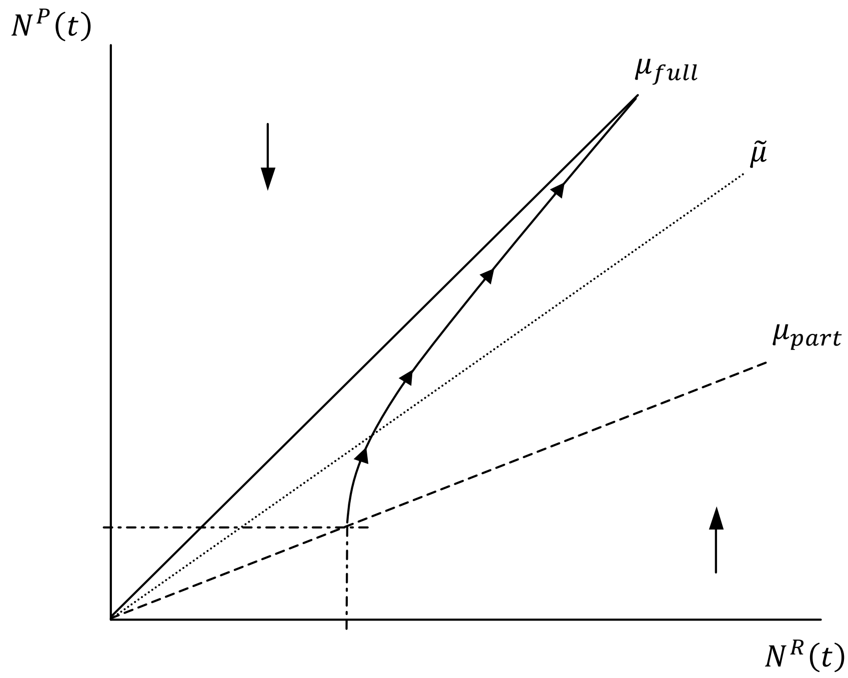


Figure 2.5: Phase diagram

2.5 Comparison to Standard Model

This section compares the baseline model to the standard model in Grossman and Helpman (1991b). First, we argue that a part-trade equilibrium cannot emerge in the standard model with CES preferences. Second, we replicate two experiments of Grossman and Helpman (1991b) regarding the diffusion of knowledge and product market integration in the context of a full-trade equilibrium.

As Grossman and Helpman (1991b) we too find that the diffusion of knowledge across countries unambiguously leads to a higher world growth rate compared to the closed economy. In particular, the increase in the growth rate relative to closed economy is only due to the diffusion of knowledge and not due to the integration of product markets. The integration of product markets still leads to the elimination of any duplicity in research efforts across countries. However, the integration of product markets compared to only the costless diffusion of knowledge across countries without product market integration has different effects. The reason is that unlike in the standard model firms cannot pass through trade costs in our model. This implies that the growth rate in the case of pure knowledge diffusion is higher than in the case of product market integration, whereas in the standard model there is no difference between product market integration and costless knowledge diffusion due to constant markups.

2.5.1 Equilibrium Structure

We compare the equilibrium structure to the standard case of CES preferences. In the standard model discussed in Grossman and Helpman (1991b) case (*ii*), a part-trade equilibrium, can never occur. The reason is the following. In the case of CES preferences the marginal utility from consuming an infinitesimal quantity of good j is infinitely high. Therefore, all households will always consume all goods available even in the presence of finite trade costs τ. In other words, the restriction in Proposition 2.1 that the price charged in the rich country cannot exceed the price charged in the poor country times trade costs is never binding. We conclude that the introduction of non-homothetic preferences of the zero-one form implies a different equilibrium structure, where the degree of inequality across countries determines the incentives to innovate and trade patterns, simultaneously.

2.5.2 Diffusion of Knowledge and Product Market Integration

Grossman and Helpman (1991b) study the question how international trade can act as an engine of growth through the diffusion of knowledge across borders. They point out three channels through which a country's relationship with other countries may affect its incentives to innovate. First, international exchange might include the exchange of technological knowledge. Second, international competition induces firms in each country to develop new products that are unique in both countries. Third, international integration might increase the market size for firms. They notice that the integration of economies has two opposing effects, which cancel each other in equilibrium. On the one hand, it increases profit opportunities by enlarging the market size of every firm. On the other hand, it reduces profit opportunities by intensifying competition. In the case of zero-one preferences there is also the effect that firms cannot pass through trade costs to consumers. In the next two sections we replicate part of Grossman and Helpman's (1991b) analysis. We focus on the case of a full-trade equilibrium to be as close as possible to the equilibrium structure in the standard model. Furthermore, we will assume symmetric technology across countries, i.e. $F^R = F^b = F$ and $b^R = b^P = b$, and denote each country's population size by L. However, we still assume $\theta_R \geq \theta_P$.

Diffusion of Knowledge

Grossman and Helpman (1991b) start their discussion with a look at the diffusion of knowledge across two countries that might differ only with respect to their endowments with efficiency units of labor. They consider opening up communication channels between countries, which enables the costless transmission of knowledge, without integrating their product markets. However, R&D that is common in both countries contributes only once to the total knowledge stock. They focus on the special case of international diffusion where the knowledge stock in both countries is given by $N^R(t) + \psi N^P(t)$, where ψ is the fraction of goods available in the poor country that is not available in the rich country. In other words, new products developed in the rich country are always unique whereas a fraction $(1 - \psi)$ of new products developed in the poor country replicates goods that are already available in the rich country. In the steady state, the labor market clearing conditions in the rich and poor country are then determined by

$$\begin{aligned} L\theta_R &= g^R \frac{F}{1+\psi\frac{N^P(t)}{N^R(t)}} + L\frac{b}{1+\psi\frac{N^P(t)}{N^R(t)}} \\ L\theta_P &= g^P \frac{F}{\frac{N^R(t)}{N^P(t)}+\psi} + L\frac{b}{\frac{N^R(t)}{N^P(t)}+\psi}. \end{aligned}$$

The labor market conditions imply that both countries grow at the same rate $g^R = g^P = g$ because in the steady state the ratio $N^R(t)/N^P(t)$ must be constant, and is determined by θ_R/θ_P. Hence, the growth rate in both countries is given by

$$g = \frac{1}{F/bL}\left(\frac{\theta_R + \psi\theta_P}{b} - 1\right) > 0 \Leftrightarrow \theta_R + \psi\theta_P > b.$$

We conclude that the diffusion of knowledge across countries unambiguously leads to a higher growth rate compared to the closed economy case where country i grows at the constant rate $g^i = (\theta_i/b - 1)/(F/bL)$. It is obvious that the smaller the overlap of research, i.e. the higher ψ, the higher the growth rate g. This is identical to Grossman and Helpman (1991b).

Product Market Integration

In the open economy equilibrium with full-trade, the production share of each country is determined by relative endowments with efficiency units of labor, i.e. $m = \theta_P/(\theta_R + \theta_P)$, and the growth rate is given by

$$g = \frac{1}{F/bL}\left(\frac{\theta_R + \theta_P}{b} - (1+\tau)\right) > 0 \Leftrightarrow \theta_R + \theta_P > (1+\tau)\, b.$$

Remember that the integration of product markets leads to competition among innovators across countries that eliminates the duplicity of new products. We observe that the growth rate g in the full-trade equilibrium is lower than in the equilibrium with international knowledge spillovers if there is no duplication of new products (i.e. $\psi = 1$). The reason is that in the case of zero-one preferences firms cannot pass through trade costs to households. Thus, profits and markups are lower in the case of product market integration compared to the case of costless knowledge diffusion. This contrasts to the case with CES preferences where firms can fully pass through trade costs. In contrast to Grossman and Helpman (1991b) there is the additional effect of the trade cost pass-through, which dominates the offsetting effects of a larger market and more intense competition.

However, the growth rate in the open economy is still larger than in the closed economy. Integrating product markets leads to higher growth compared to the closed economy if and only if $\theta_R \geq \theta_P > \tau b$, in which case we have $g > g^R \geq g^P$. The condition $\theta_P > \tau b$ must hold in a trade equilibrium. A trade equilibrium exists, if and only if each firm's profits from trading exceed its profits from not trading. One can show that for a firm located in the rich country profits under trade exceed those under autarky if and only if $\theta_P > \tau b$. Since $\theta_R \geq \theta_P$ the condition automatically holds for firms located in the poor country. Intuitively, the increase in market size must, ceteris paribus, more than compensate for the decrease in profits per unit sold brought about by the increase in marginal cost through trade cost, so that profits of each firm and therefore innovation incentives increase.

2.6 Applications and Extensions

In this section we discuss applications in the form of policy changes and their welfare consequences as well as extensions like alternate assumptions about knowledge spillovers and preferences. Instead of working out general cases we concentrate on some special cases that highlight interesting points. All of the issues addressed in this section arise naturally through the introduction of non-homothetic preferences.

2.6.1 Intellectual Property Rights Policy

We use our model to study how the design of intellectual property rights determines the incentives to innovate and trade patterns. So far, we have assumed that there is international exhaustion of patents, in other words, parallel imports cannot be legally prevented. Now suppose that policymakers in both countries simultaneously choose to introduce national exhaustion of patents so that parallel imports can be legally prevented.[4] In that case the price setting restriction is never binding regardless of the level of inequality across countries or trade

[4]Suppose that the rich country does not allow parallel imports but the poor country does. This implies that the price setting restriction is never binding, and all firms located in the rich country always export. Consider the opposite case where the rich country allows parallel imports but the poor country does not. In that case, some firms in the rich country might have an incentive not to export. Note that in both cases, the policy of the poor country does not affect the incentives to export of firms in the rich country.

costs. Hence, the economy is always in a full-trade equilibrium. Since monopolistic firms can fully price discriminate across countries it is never a dominant strategy for a firm located in the rich country to sell exclusively on its domestic market as long as households in the poor country are not too poor, i.e. $z_P \geq \tau w^R b^R$. This implies that more resources are absorbed in production, relative to a part-trade equilibrium under international exhaustion, so that less resources can be allocated to research and development. We conclude that national exhaustion of patents would depress growth but would encourage international trade at the extensive margin. Households in the poor country suffer a dynamic loss as incomes grow at a lower g but experience a static gain because they consume a larger share n of products, whereas households in the rich country only suffer a dynamic loss. Since households in the rich country suffer only losses they would almost surely favor international exhaustion of patents. Whether households in the poor country favor international or national exhaustion depends on their time preference rate, i.e. how much they discount future losses against current gains.

The design of property rights and its effects on the incentives to innovate have been a hot topic for some time in the public (e.g. trade-related aspects of intellectual property rights TRIPS) and academic debate. This is especially true for pharmaceutical products in the US where prices of identical drugs differ substantially from those in Canada or Mexico. A recent example from the academic literature is Grossman and Lai (2008) who find that the innovation rate is faster in a world with international exhaustion of patents compared to national exhaustion. In their model, parallel imports are induced by different national price controls, in contrast to our model where the extent of income inequality across countries induces parallel imports. Grossman and Lai (2008) argue that a switch from national to international exhaustion of patents changes innovation incentives because the government follows different motives in its price regulation with and without parallel imports.

2.6.2 Trade Policy

Our model can be applied to study the effects of trade liberalization in the form of a reduction of tariffs induced for example by the General Agreement on Tariffs and Trade (GATT). Suppose that such a policy liberalizing trade is put in place. We can interpret a decrease in trade costs $\tau > 1$ as a reduction in tariffs. Note that τ captures all possible barriers to trade in goods, e.g. efficiency of the transportation sector, or non-tariff barriers to trade. Note that with zero-one preferences trade costs are borne by firms whereas with CES preferences firms pass them on to consumers. A decrease in trade costs increases profits and therefore the incentives to innovate, ceteris paribus. If the fall in trade costs implies that the price setting restriction for firms becomes binding the economy ends up in a part-trade steady state. We know that, ceteris paribus, in a part-trade steady state the incentives to innovate are higher than in a full-trade steady state. The equivalent argument is that a decrease in trade costs releases resources from the production sector that can be allocated to research and development (in both countries). If trade costs fall sufficiently the economy might end up in a part-trade equilibrium. The intuition is that at low trade costs the price setting restriction may be binding for firms, for a given degree of inequality across countries. In that case, further resources are released from the production sector since households in the poor country can no longer afford to consume all goods produced in the rich country. Either way, the growth rate of the global economy increases if trade costs fall. We conclude that trade liberalization has a positive effect on the growth rate and might have a negative effect on the extensive margin of trade. Our results are in line with the empirical evidence, which suggests a negative relationship between trade restrictions and economic growth. A critical survey of the empirical literature can be found in Rodriguez and Rodrik (2001). In contrast, in the standard model with CES preferences trade costs have no effect on the incentives to innovate because they are passed on to consumers (see additionally Feenstra 2004).

2.6.3 Welfare

We focus our welfare analysis on the effect of a trade liberalization on relative welfare levels. The advantage of that analysis is that we can shut down transfers, i.e. $T(t) = 0$ for all t, and trace the transitional dynamics analytically. First, in equilibrium the difference in welfare levels between households in the rich and the poor country, i.e. $\Delta V(0) \equiv V_R(0) - V_P(0)$, is given by

$$\Delta V(0) = \int_0^\infty e^{-\rho t} \log\left(\frac{N_R(t)}{N_P(t)}\right) dt.$$

Remember, in a full-trade equilibrium $N_R(t) = N_P(t) = N(t)$ for all t, whereas in a part trade equilibrium $N_P(t) < N_R(t)$ for all t. In order to analyze the effect of a change in trade policy on welfare consider a decrease of τ, as discussed in Section 2.6.2, from a level that sustains a full-trade equilibrium to one that sustains a part-trade equilibrium. First, note that there is a critical $\widetilde{\mu} > 0$, where $(1-\beta)\widetilde{\mu}/\beta \equiv \tau_{crit}$ and $\tau_{crit} \equiv z_R/z_P$, at which point the economy switches from a full- to a part-trade equilibrium. First, in a steady state with high trade costs τ we have a relatively high $\mu > \widetilde{\mu}$ whereas in a steady state with low trade costs $\mu < \widetilde{\mu}$ is also low. In a part-trade equilibrium relatively more firms enter in the rich than in the poor country because in the rich country resources are released from the production sector that can be allocated to the research sector due to the fact that households in the poor country can no longer afford to purchase all products produced in the rich country. Hence, during the transition $\mu(t)$ falls from a level above $\widetilde{\mu}$ to one below. As long as $\mu(t) > \widetilde{\mu}$, the economy is in a full-trade equilibrium, and $\Delta V(0)$ is equal to zero since $N_R(t) = N_P(t)$. Eventually, as $\mu(t) < \widetilde{\mu}$ the economy switches to a part-trade equilibrium, and the difference in welfare can be written as (using the balance of payments condition in the part-trade equilibrium)

$$\Delta V(0) = \int_0^\infty e^{-\rho t} \log B\left(1 + \frac{1}{\mu(t)}\right) dt > 0$$

where $B \equiv \beta/[\beta + \tau(1-\beta)]$. We know that $\mu(t)$ falls until it has reached its new steady state value $\mu < \widetilde{\mu}$. This implies that $\Delta V(0)$ becomes positive as the economy switches from a full- to a part-trade equilibrium, increases during the transition until the economy has reached the new steady state, and stays constant at a positive value from then on. Hence, we conclude that households in the poor country unambiguously loose relative to households in the rich country from a trade liberalization in the form of a decrease in trade costs τ. Intuitively, households across countries share a common growth rate in and out of the steady state of their corresponding consumption sets regardless of whether the economy is in a full- or part-trade equilibrium, and differ only with respect to the sizes of their consumption sets N_i if the economy is in a part-trade equilibrium.[5] Therefore, households in the poor country only experience a relative loss in the sense that the set of goods they consume $N_P(t)$ falls relative to the set of goods consumed by households in the rich country $N_R(t)$ as the economy moves from a full- to a part-trade equilibrium. In sum, households in the poor country might be opposed to a trade liberalization if trade costs decrease so much that firms can no longer perfectly price discriminate across countries and some firms in the rich country decide to no longer serve them.

2.6.4 Knowledge Spillovers: Learning-by-importing

Remember that Coe and Helpman's (1995) estimates suggest that foreign R&D has a more beneficial effect on domestic productivity the more open an economy is to foreign trade. We pick this idea up and look at the case when intertemporal knowledge spillovers in research and development only originate from domestically produced goods and imported goods. In our model the extent of international knowledge spillovers is endogenously determined by trade patterns, which in turn are determined by income inequality across countries. Technology in country $i = \{R, P\}$ is now as follows. To develop a new product $F^i(t) = F/N_i(t)$ units of labor are required, and to produce 1 unit of output it needs $b(t) = b/N(t)$ units of labor. For simplicity we assume that technology is symmetric across countries except for the extent of international knowledge spillovers in R&D. Furthermore, we assume that each country's population size is given by L. First, suppose that inequality across countries is low for a given level of trade costs so that the economy is in a full-trade equilibrium, i.e. $N_R(t) = N_P(t) = N(t)$. Both countries import all goods produced in the other country so that there are perfect international knowledge spillovers. In that case, the world growth rate in the steady state is given by

$$g = \frac{1}{F/bL}\left(\frac{\theta_R + \theta_P}{b} - (1+\tau)\right) > 0 \;\Leftrightarrow\; \theta_R + \theta_P > (1+\tau)b.$$

Now, consider the case where inequality across countries is high (or trade costs are low) so that the economy is in a part-trade equilibrium. Hence, $N_P(t) = N^P(t) + N_M^R(t) < N_R(t) = N(t)$ so that $N_P(t)/N(t) = n < 1$. To solve for the steady state, we only have to modify the resource constraint of the poor country as follows

$$\beta L \theta_P = g\left(\frac{m}{n}\right) F + m\,(1+\tau)\, bL.$$

[5]However, note that consumption expenditures evolve differently across countries during the transition. Since $\mu = N^P/N^R$ falls over time it must be that aggregate consumption expenditures in the rich country grow at a lower rate than in the poor country, i.e. investment in the rich country exceeds investment in the poor country.

Due to our choice of numeraire, i.e. $w^P(t)b^P(t) = 1$, the rest of the equations for the part-trade equilibrium in Appendix 2.A.2 continue to apply. First, note that if there are no transfers $T = 0$ the growth rate is determined by the resource constraints in the poor and rich countries. In the general case with $T \neq 0$ one can show with the help of numerical simulations that an increase in inequality across countries through a regressive transfer increases the steady state growth rate g. Intuitively, if households in the poor country become even poorer they can afford to purchase less products as before. Hence, labor in the R&D sector of the poor country becomes, ceteris paribus, less productive, which increases the cost of innovation in the poor country. At the same time, households in the rich country become even richer, which increases the incentives to innovate in the rich country. In our simulations, the positive effect on growth in the rich country still dominates so that the global growth rate g increases.

Limit Case: No Learning-by-importing

If there are no international knowledge spillovers, there might be uneven growth across countries. In the full-trade steady state, the rich, respectively, poor country grows at the rate

$$\begin{aligned} g^R &= \frac{1}{F^R/b^R L}\left(\frac{(1-\beta)\theta_R}{b^R} - [(1-\beta)+\tau\beta]\right) \Leftrightarrow (1-\beta)\theta_R > [(1-\beta)+\tau\beta]b^R \\ g^P &= \frac{1}{F^P/b^P L}\left(\frac{\beta\theta_P}{b^P} - [\beta+\tau(1-\beta)]\right) \Leftrightarrow \beta\theta_P > [\beta+\tau(1-\beta)]b^P. \end{aligned}$$

In the case, where technology is symmetric, $g^R > g^P$ if and only if $(1-\beta)\theta_R - \beta\theta_P > 1-\tau$. This implies that if all households have the same endowments with efficiency units of labor $\theta_R = \theta_P = \theta$ a sufficient condition for $g^R > g^P$ is $\beta < 0.5$. In other words, the relatively larger country in terms of population size grows at the higher rate due to the home market effect. Furthermore, if $\beta = 0.5$ a necessary and sufficient condition for $g^R > g^P$ is $\theta_R > \theta_P$. Intuitively, the country with the relatively higher labor endowment grows at the higher rate due to the increasing-returns-to-scale technology. We see that the distribution of income across countries still has no effect on the incentives to innovate. In the part-trade steady state, the balance of payments requires that $m = N^P(t)/N(t)$ is constant. However, this implies that the growth rate in both countries must be the same, i.e. $g^R = g^P = g$. Hence, the world growth rate g is determined by the resource constraint in the poor country

$$g = \frac{1}{F^P/b^P L}\left(\frac{\beta\theta_P}{b^P} - [\beta+\tau(1-\beta)]\right) \Leftrightarrow \beta\theta_P > [\beta+\tau(1-\beta)]b^P.$$

The resource constraint in the rich country $(1-\beta)L\theta_R = gF^R + b^R(1-\beta)L + q\tau b^R\beta L$ then pins down the number of firms located in the rich country that export q. In other words, the number of exporting firms in the rich country adjusts such that given the world growth rate determined in the poor country, the labor market in the rich country clears. Apparently, a regressive transfer, which increases inequality, has now no effect on the growth rate g. A regressive transfer only results in an adjustment of the balance of payments, i.e. the willingness to pay of households in the poor country z_P decreases and/or the production share of the poor country m increases, such that q remains consistent with labor market clearing in the rich country. We conclude that in the part-trade steady state the incentives to innovate are determined by technology and endowments in the two countries and are independent of the distribution of income across countries if there are no international knowledge spillovers. However, inequality across countries still has an effect on innovation incentives and trade in the sense that if inequality increases sufficiently the economy switches from a full- to a part-trade steady state. This implies that a sufficiently high increase in inequality might now lead to a decrease in the average growth rate if and only if $g^R > g^P$. We further get convergence in growth rates but divergence in consumption levels.

2.6.5 Quadratic Preferences

We briefly discuss the case where households can adjust their consumption along the extensive and intensive margin. The households' quasi-homothetic utility function takes the following form

$$u\left(\{c(j,t)\}_{j=0}^{\infty}\right) = \int_{j=0}^{\infty} \left(sc\,(j,t) - \frac{1}{2}c\,(j,t)^2\right) dj \tag{2.42}$$

where $s > 0$ denotes a local saturation level. The utility function has the following (standard) properties, $u'(\cdot) = s - c\,(j,t) > 0$, $u''\,(c\,(j,t)) = -1 < 0$, and $\lim_{c(j,t)\to 0} u'(\cdot) = s < \infty$, which implies that marginal utility is bounded from above. Therefore, non-negativity constraints $c\,(j,t) \geq 0$ might become binding for some j's. Thus, households in the poor country might not be able to afford all products made in the rich country. It is straightforward to show that if income inequality across countries is sufficiently high the price setting restriction imposed by the threat of parallel imports becomes binding for firms located in the rich country. Some firms producing in the rich country don't export to the poor country. Hence, a part-trade equilibrium can emerge even if households have not only a choice about how many different goods they consume but also about how much of each good they consume.

We restrict our discussion to the effect of income inequality across countries on innovation incentives and trade patterns. We make the same experiment as in Section 2.4.1, and simulate a change in income inequality across countries by making regressive transfers, starting out in a full-trade steady state, and ending up in a part-trade steady state. The equations describing the steady state can be found in Appendix 2.A.6. The results are summarized as follows: An increase in inequality through a regressive transfer has a positive effect on the growth rate, the terms of trade for the poor country, and a negative effect on the production share of the poor country, the share of exporting firms, and the consumption share of households in the poor country.

We compare these results with the baseline model. Since the results and their intuition are similar, we keep the discussion relatively short. Given that the rich consume all goods available they can only increase their consumption by increasing the amount of each good they purchase. However, since they are satiated at one point, the price elasticity of demand for those goods decreases. Hence, markups rise, and the incentive to innovate tends to increase. At the same time, households in the poor country become poorer and decrease the consumption of each good they purchase, ceteris paribus. The price elasticity of demand tends to increase, markups to fall, so that the incentive to innovate decreases. The economy is in a full-trade equilibrium as long as households in the poor country become not too poor such that they also decrease the number of goods they purchase relative to the number of goods that are available on the market, else the economy is in a part-trade equilibrium. In both cases, simulations show that the effect of the households in the rich country on the incentives to innovate dominate, so that we see more innovation in a steady state with higher inequality across countries. This effect is even stronger in a part-trade equilibrium when households in the rich country are very rich, and pay very high markups. The intuition for the movement of the terms of trade in favor of the poor country, for the decrease in the share of exporting firms in the rich country, and the fall in the production share of the poor country is the same as in our baseline model (see discussion in Section 2.4.1).

It is worth noting two things that are different to the baseline model. First, since there is an intensive margin of consumption the incentives to innovate in the full-trade equilibrium change as inequality across countries changes. Second, the extensive and intensive margin of trade is higher in a full-trade equilibrium than in a part-trade equilibrium.

In sum, the introduction of quadratic preferences does not change the basic results and intuition of the baseline model. The more similar countries are the lower the incentives to innovate, and the higher the extensive and intensive margins of trade.

2.7 Conclusion

We look at the consequences of inequality in per capita incomes across countries on the incentives to innovate in the global economy and the patterns of trade between countries. To this end, we introduce non-homothetic preferences in the standard model of Grossman and Helpman (1991b). We show that if income inequality across countries is high, some firms located in the rich country have an incentive to sell exclusively in the domestic market. In order to circumvent the threat of parallel imports and be able to charge high prices they choose to forgo a larger market. At the same time, this means that not all goods produced in the world economy are traded. Since households in the poor country cannot afford to buy all goods produced in the rich country, relatively more resources in the rich country can be allocated to research and development. In sum, at high levels of inequality across countries the incentives to innovate are high whereas the extensive margin of trade is low.

We apply the model to various questions that arise naturally in the international context, and discuss several extensions. First, we apply the model to the question of intellectual property rights in an international context. We show that households in the rich and poor country might not see eye to eye about the design of intellectual property rights. Whether they agree or disagree depends crucially on how much households in the poor country weigh future losses in consumption against present gains. Second, we take a closer look at trade policy, and argue that a reduction in trade costs might increase the incentives to innovate but decrease the extensive margin of trade. Such a change in trade policy might make households in the poor country unambiguously worse off relative to households in the rich country if trade costs fall sufficiently such that the economy moves from a full- to a part-trade equilibrium. We extend the model by making the degree of international knowledge spillovers endogenous. In particular, we assume that knowledge spillovers in the research sector depend on the set of products consumed in a given country. This implies that firms benefit only from knowledge created abroad if these products are imported. We show that the results from the baseline model are robust, i.e. inequality across countries still leads to an increase in the growth rate and a decrease in the extensive margin of trade. We also consider an extreme case by shutting down international knowledge spillovers completely. We show that countries grow at different rates in a full-trade equilibrium, and converge to a common growth rate, which is equal to the growth rate of the poor country, in a part-trade equilibrium. In this limit case an increase in inequality might lead to a decrease in the average growth rate in the economy as well as the extensive margin of trade. Last, we show that the relationship between inequality, trade patterns, and the growth rate is robust to the introduction of preferences where households have a choice along the extensive and intensive margin of consumption.

2.A Appendix

2.A.1 Full-trade Equilibrium

Balance of Payments

Note that due to Walras' law the balance of payments is implied by the budget constraints, labor market clearing conditions, and the zero-profit conditions.

Let us start with the aggregate flow budget constraint in the poor country in the full-trade equilibrium

$$\dot{A}_P(t) = r^P(t)A_P(t) + w^P(t)\beta L\theta_P - \beta LN(t)z_P(t) + \beta LT_P(t).$$

If we choose the marginal cost of production for firms located in the poor country as numeraire, i.e. $w^P(t)b^P(t) = 1$, we can write the flow budget constraint as follows

$$\frac{\dot{N}^P(t)}{N^P(t)}\frac{F^P}{b^P} = r^P(t)\frac{F^P}{b^P} + \frac{N(t)}{N^P(t)}\frac{\beta L\theta_P}{b^P} - \beta L\frac{N(t)}{N^P(t)}z_P(t) + \frac{N(t)}{N^P(t)}\frac{\beta LT_P(t)}{N(t)}$$

where capital market clearing in country P implies that $A_P(t) = N^P(t)F^P/b^P$.

Next, the labor market clearing condition in the poor country is determined by (2.29) and can be written as

$$\frac{\dot{N}^P(t)}{N^P(t)}\frac{F^P}{b^P} = \frac{N(t)}{N^P(t)}\frac{\beta L\theta_P}{b^P} - \beta L - \tau(1-\beta)L.$$

We can rewrite the zero-profit condition in the poor country as

$$r^P(t)\frac{F^P}{b^P} = [z_P(t) - 1]\,\beta L + [z_R(t) - \tau]\,(1-\beta)L.$$

Substituting the labor market clearing and the zero-profit condition into the flow budget constraint yields the balance of payments

$$\begin{aligned}
\frac{\beta L\theta_P}{m(t)b^P} - \beta L - \tau(1-\beta)L &= [z_P(t) - 1]\,\beta L + [z_R(t) - \tau]\,(1-\beta)L \\
&+ \frac{\beta L\theta_P}{m(t)b^P} - \frac{\beta Lz_P}{m(t)} + \frac{1}{m(t)}\frac{\beta LT_P(t)}{N(t)} \\
[1 - m(t)]\,\beta Lz_P(t) &= m(t)(1-\beta)Lz_R(t) + \beta L\frac{T_P(t)}{N(t)} \\
N^R(t)\beta Lz_P(t) &= N^P(t)(1-\beta)Lz_R(t) + \beta LT_P(t)
\end{aligned}$$

where $m(t) \equiv N^P(t)/N(t)$. Alternatively, we could have derived the balance of payments from the aggregate flow budget constraint, the labor market clearing condition, and the zero-profit condition in the rich country.

Steady State

The following equations determine the full-trade steady state

$$\begin{aligned}
g &= r-\rho \\
(1-\beta)\,L\theta_R &= g\,(1-m)\,F^R+(1-m)\,b^R\,(1-\beta)\,L+(1-m)\,\tau b^R\beta L \\
\beta L\theta_P &= gmF^P+mb^P\beta L+m\tau b^P\,(1-\beta)\,L \\
\omega^R F^R &= \frac{\left(z_R-\omega^R b^R\right)(1-\beta)L+\left(z_P-\tau\omega^R b^R\right)\beta L}{r} \\
\frac{F^P}{b^P} &= \frac{(z_P-1)\,\beta L+(z_R-\tau)\,(1-\beta)L}{r} \\
(1-m)\,\beta Lz_P &= m(1-\beta)Lz_R+\beta LT
\end{aligned}$$

where $\omega^R \equiv w^R(t)/N(t)$ and $T \equiv T_P(t)/N(t)$.

Prices and Relative Wages in Steady State

We can solve for prices z_R and z_P as functions of the endogenous variables g and m as follows

$$\begin{aligned}
z_R &= \frac{1-m}{1-\beta}\left\{\frac{F^P}{b^P L}(g+\rho)+[\beta+\tau(1-\beta)]\right\}-\frac{\beta}{1-\beta}T \\
z_P &= \frac{m}{\beta}\left\{\frac{F^P}{b^P L}(g+\rho)+[\beta+\tau(1-\beta)]\right\}+T.
\end{aligned}$$

We see that there is a positive relationship between prices (or the willingness to pay) and the growth rate g for a given production share of the poor country m. Furthermore, the relationship between z_R and m is negative, and between z_P and m positive, ceteris paribus. For the case of identical technologies across countries, i.e. $F^R = F^P = F$ and $b^R = b^P = b$, we can formulate the following proposition

Proposition 2.5. *The larger country has ceteris paribus the higher wage rate.*

Proof. In general, we can write the (relative) wage rate ω^R in the rich country as follows

$$\omega^R b^R = \frac{\frac{F^P}{b^P}\left[(1-\beta)z_R+\beta z_P\right]}{\frac{F^R}{b^R}\left[(1-\beta)z_R+\beta z_P\right]+\frac{F^P}{b^P}\left[(1-\beta)+\tau\beta\right]-\frac{F^R}{b^R}\left[\beta+\tau(1-\beta)\right]}.$$

Imposing symmetric technologies across countries, we can write

$$\omega^R b = \frac{\left[(1-\beta)z_R+\beta z_P\right]}{\left[(1-\beta)z_R+\beta z_P\right]+(1-\tau)(1-2\beta)}.$$

It is straightforward to see that if technologies and population sizes are identical, i.e. $\beta = 1/2$, across countries wage rates equalize, i.e. $\omega^R b = 1$ (remember our choice of numeraire $\omega^P b = 1$). If $\beta < 1/2$ the second term in the denominator is negative (remember that $\tau > 1$) so that the numerator is larger than the denominator, and therefore $\omega^R b^R > 1$. In other words, as in the standard model with CES preferences the larger country has the higher wage rate as a result of the home market effect (see e.g. Krugman 1980). If $\tau = 1$ there is no home market effect, and we see that wages equalize across countries. □

2.A.2 Part-trade equilibrium

Steady State

The following equations determine the part-trade steady state

$$\begin{aligned}
g &= r-\rho \\
(1-\beta)L\theta_R &= g\,(1-m)\,F^R+(1-m)\,b^R(1-\beta)L+q\,(1-m)\,\tau b^R\beta L \\
\beta L\theta_P &= gmF^P+mb^P\beta L+m\tau b^P(1-\beta)L \\
\left(z_R-\omega^R b^R\right)(1-\beta)L &= \left(\tau z_P-\omega^R b^R\right)(1-\beta)L+\left(z_P-\tau\omega^R b^R\right)\beta L \\
\omega^R F^R &= \frac{\left(z_R-\omega^R b^R\right)(1-\beta)L}{r} \\
\frac{F^P}{b^P} &= \frac{(z_P-1)\,\beta L+(\tau z_P-\tau)\,(1-\beta)L}{r} \\
\beta Lq\,(1-m)\,z_P &= (1-\beta)Lm\tau z_P+\beta LT
\end{aligned}$$

where $\omega^R \equiv w^R(t)/N(t)$ and $T \equiv T_P(t)/N(t)$.

Prices and relative wages in steady state

We can solve for prices as a function of the endogenous variable g as follows

$$\begin{aligned}
z_R &= \left[1+\frac{g+\rho}{(1-\beta)L}\right]\omega^R b^R \\
z_P &= 1+\frac{F^P}{b^P L}\frac{g+\rho}{[\beta+\tau(1-\beta)]}
\end{aligned}$$

where the wage rate $\omega^R b^R$ is determined below. As in the full-trade steady state, if technologies are identical across countries, we propose the following:

Proposition 2.6. *The larger country has the higher wage rate, holding everything else constant.*

Proof. In general, we can write the (relative) wage rate ω^R as follows

$$\omega^R b^R = \frac{\frac{F^P}{b^P}(g+\rho)+[\beta+\tau(1-\beta)]L}{\frac{F^R}{b^R}(g+\rho)+[(1-\beta)+\tau\beta]L}.$$

Imposing symmetric technologies across countries implies that

$$\omega^R b = \frac{\frac{F}{b}(g+\rho)+[\beta+\tau(1-\beta)]L}{\frac{F}{b}(g+\rho)+[(1-\beta)+\tau\beta]L}.$$

It is easy to see that if population sizes are identical in both countries, i.e. $\beta = 1/2$, wage rates equalize, i.e. $\omega^R b = 1$ (due to our choice of numeraire). If $\beta < 1/2$ it is straightforward to show that $\omega^R b > 1$, and vice versa for $\beta > 1/2$. This is identical to the full-trade equilibrium. □

2.A.3 Inequality Simulations

To simulate the effects of a change in between-country income inequality we choose the following parameter values $L = 1$, $\beta = 0.5$, $\theta_R = 2.5$, $\theta_P = 2$, $F^R = F^P = 7$, $b^R = b^P = 1$, $\rho = 0.04$, $\tau = 1.6$, and $T \in (-0.65, 0.15)$ varying. Note that since technologies and population sizes are identical across countries wage rates equalize. We assume identical technology and population size across countries to isolate the effects of inequality. We measure inequality with the Gini coefficient, which we construct from the Lorenz curve (see Ray 1998). Note, that relative lifetime incomes in per capita terms can be written in the steady state as

$$\frac{y_R(t)}{y_P(t)} = \frac{w^R(t)\theta_R + \rho a_R(t) - T_R(t)}{w^P(t)\theta_P + \rho a_P(t) + T_P(t)}.$$

Obviously, if the poor country's income share equals its population share β the Lorenz curve lies on the 45 degree line of perfect equality, and the Gini coefficient is zero.

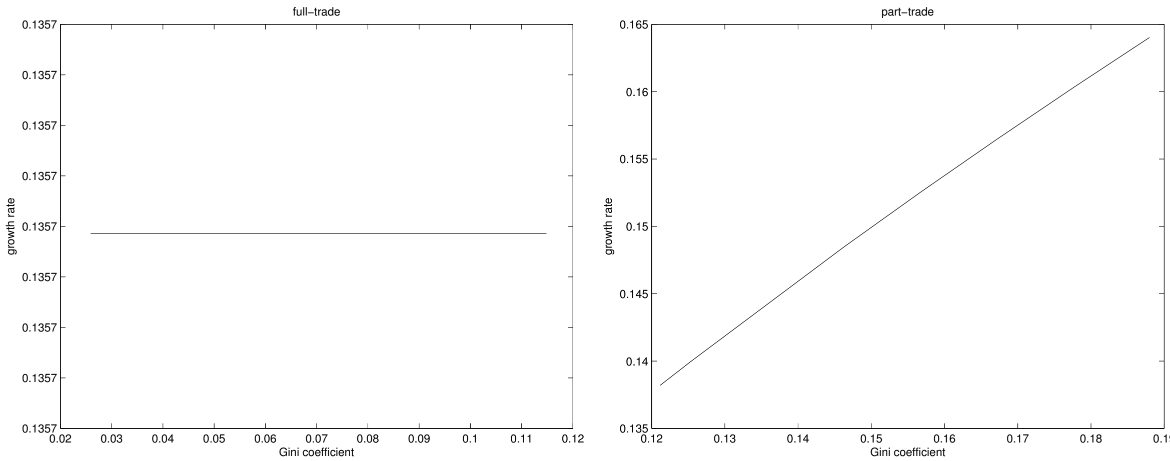


Figure A.1: Effect of inequality on growth rate in full-trade equilibrium (left-hand side) and in part-trade equilibrium (right-hand side)

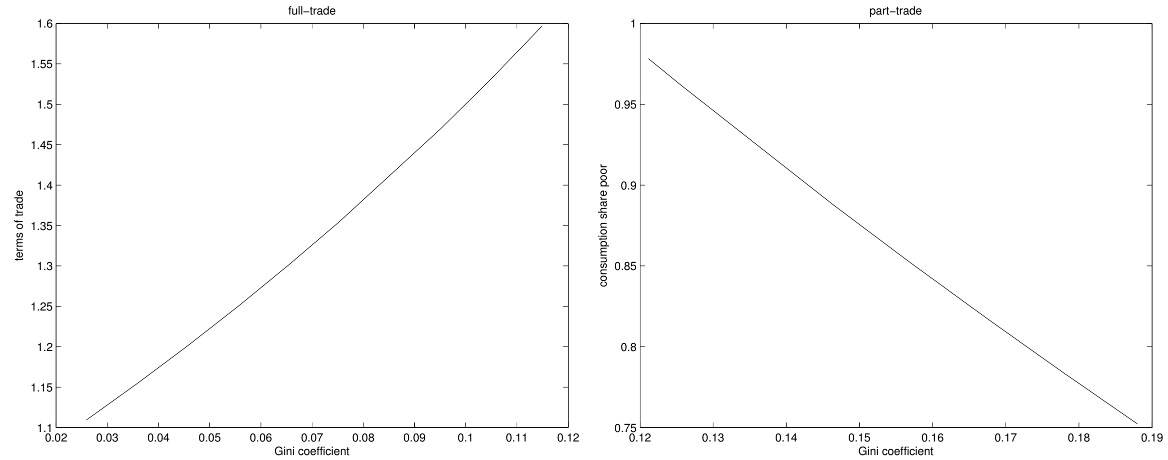


Figure A.2: Effect of inequality on terms of trade in full-trade equilibrium (left-hand side), and on relative consumption poor to rich n in part-trade equilibrium (right-hand side)

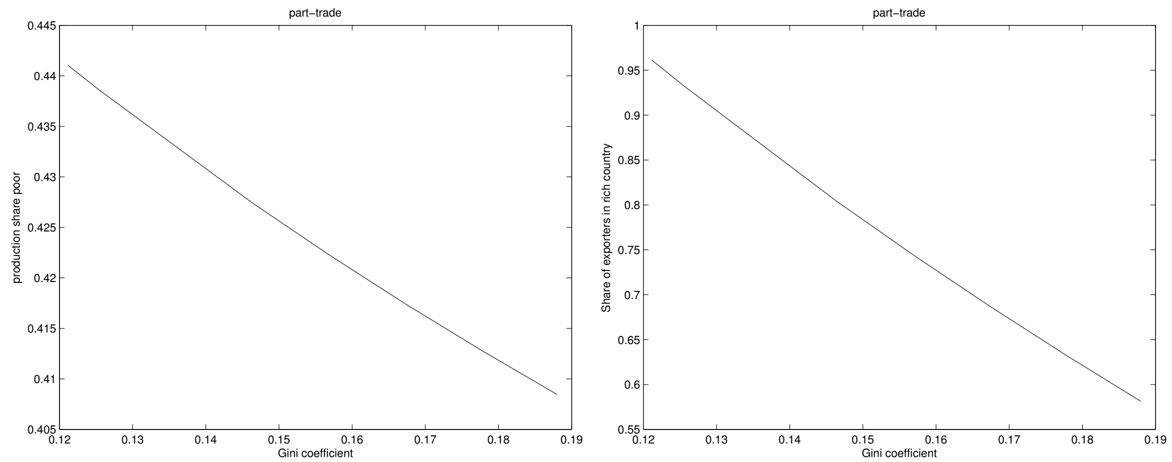


Figure A.3: Effect of inequality on production share m in part-trade equilibrium (left-hand side), and on the share of exporting firms q in part-trade equilibrium (right-hand side)

2.A.4 Technology Simulations

We simulate an increase in relative labor productivity in R&D in the poor country by choosing the following parameter values $L = 1$, $\beta = 0.5$, $\theta_R = 2.5$, $\theta_P = 2$, $F^R = 10$, $b^R = 1$, $b^P = 1$, $\rho = 0.04$, $\tau = 1.6$, $T = 0$, and $F^P \in (11, 25)$ varying. Note that since technologies are no longer identical across countries wage rates do not equalize.

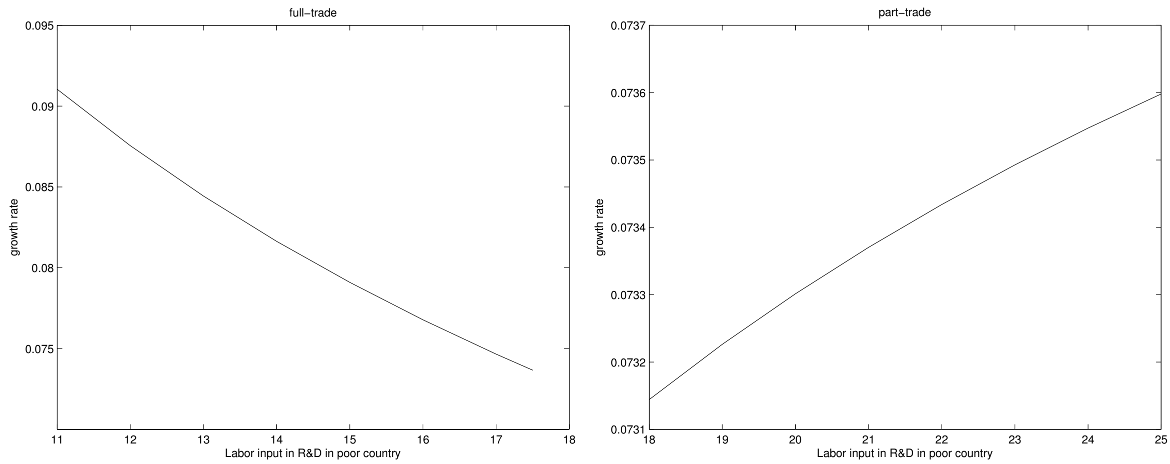


Figure A.4: Effect of labor productivity in R&D on growth rate in full-trade equilibrium (left-hand side) and in part-trade equilibrium (right-hand side)

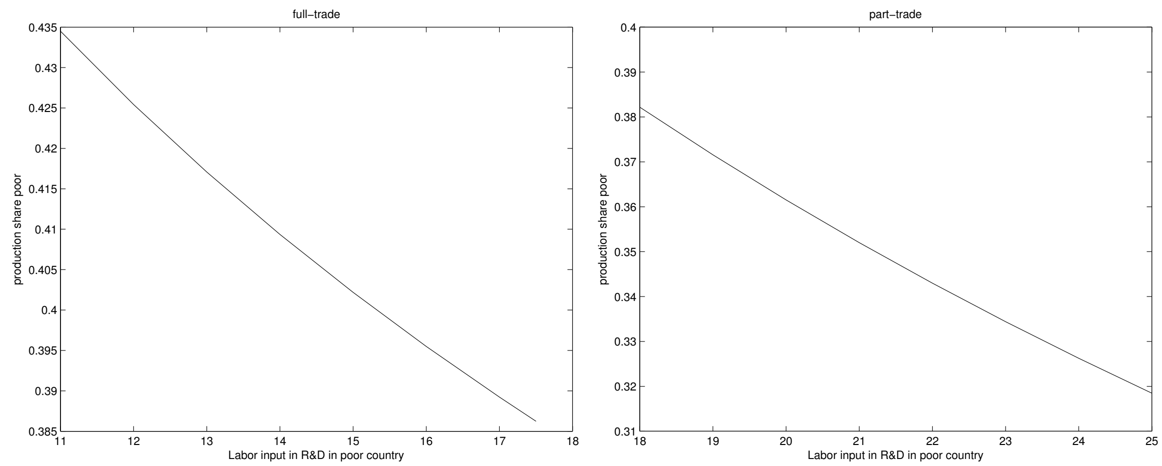


Figure A.5: Effect of labor productivity in R&D on production share of poor country in full-trade equilibrium (left-hand side) and in part-trade equilibrium (right-hand side)

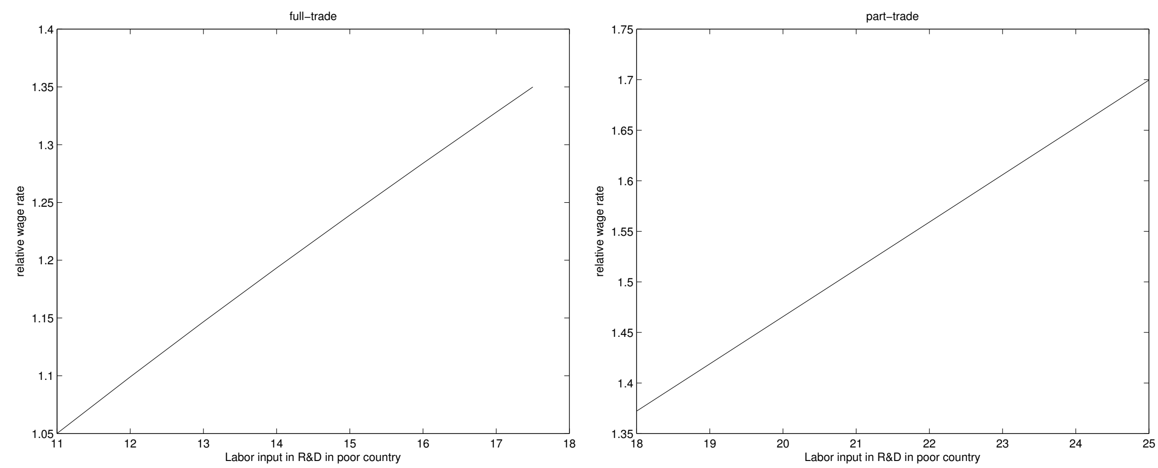


Figure A.6: Effect of labor productivity in R&D on relative wages in full-trade equilibrium (left-hand side) and in part-trade equilibrium (right-hand side)

2.A.5 Transition

It is straightforward to solve the system of linear differential equations given in the text (see e.g. Acemoglu 2008). We write the system in matrix form and define the 2×2 matrix Q as follows (for ease of notation we drop the subscript k)

$$Q \equiv \begin{bmatrix} \Phi^R & \Psi^R \\ \Psi^P & \Phi^P \end{bmatrix}$$

Note that the constants Φ_k^R, Ψ_k^R, Φ_k^P and Ψ_k^P are defined as follows. In the full-trade equilibrium, we have $\Phi_{full}^R \equiv \left[(1-\beta)L\theta_R - [(1-\beta)+\tau\beta]\, b^R L\right]/F^R$, $\Psi_{full}^R \equiv (1-\beta)L\theta_R/F^R > 0$, $\Phi_{full}^P \equiv \left[\beta L\theta_P - [\beta+\tau(1-\beta)]\, b^P L\right]/F^P$, and $\Psi_{full}^P \equiv \beta L\theta_P/F^P > 0$. Whereas in the part-trade equilibrium, we have $\Phi_{part}^R \equiv (1-\beta)L\left(\theta_R - b^R\right)/F^R > 0$, $\Psi_{part}^R \equiv (1-\beta)L\left(\theta_R - \tau^2 b^R\right)/F^R > 0$, $\Phi_{part}^P \equiv \left[\beta L\theta_P - [\beta+\tau(1-\beta)]\, b^P L\right]/F^P$, and $\Psi_{part}^P \equiv \beta L\theta_P/F^P > 0$, where $\Phi_{part}^R > 0$ and $\Psi_{part}^R > 0$ follows from Proposition 2.4.

The matrix Q can be decomposed as follows, $Q = P\zeta P^{-1}$, where P is a matrix whose columns correspond to the eigenvectors of Q and ζ is a diagonal matrix with the eigenvalues of Q on its diagonal. By the definition of an eigenvalue, $\det(Q - \zeta I) = 0$, the eigenvalues of Q are determined by the following quadratic equation

$$\zeta^2 - \left(\Phi^R + \Phi^P\right)\zeta + \left(\Phi^R\Phi^P - \Psi^R\Psi^P\right) = 0$$

with the solutions given by

$$\zeta_{1,2} = \frac{\left(\Phi^R + \Phi^P\right) \pm \sqrt{\left(\Phi^R + \Phi^P\right)^2 - 4\left(\Phi^R\Phi^P - \Psi^R\Psi^P\right)}}{2}$$

By the definition of an eigenvector, $(Q - \zeta_1 I)\, v^1 = 0$, the eigenvector v^1 associated with the eigenvalue ζ_1 is determined by following system of linearly dependent equations

$$\begin{bmatrix} \left(\Phi^R - \zeta_1\right) v_1^1 + \Psi^R v_2^1 \\ \Psi^P v_1^1 + \left(\Phi^P - \zeta_1\right) v_2^1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

By normalizing $v_2^1 = 1$, respectively, $v_2^2 = 1$, it follows that the eigenvector v^1, respectively v^2, is given by

$$v^1 = \begin{bmatrix} \frac{\zeta_1 - \Phi^P}{\Psi^P} \\ 1 \end{bmatrix}, \quad v^2 = \begin{bmatrix} \frac{\zeta_2 - \Phi^P}{\Psi^P} \\ 1 \end{bmatrix}$$

Hence, the matrices P and ζ are

$$P = \begin{bmatrix} v_1^1 & v_1^2 \\ v_2^1 & v_2^2 \end{bmatrix} = \begin{bmatrix} v_1^1 & v_1^2 \\ 1 & 1 \end{bmatrix}, \quad \zeta = \begin{bmatrix} \zeta_1 & 0 \\ 0 & \zeta_2 \end{bmatrix}$$

The system of equations is decoupled by canonical variables decomposition. The decoupled (transformed) system is easily solved, and retransformed. The solution is then given by

$$\begin{bmatrix} N^R(t) \\ N^P(t) \end{bmatrix} = \begin{bmatrix} v_1^1 c_1 \exp(\zeta_1 t) + v_1^2 c_2 \exp(\zeta_2 t) \\ c_1 \exp(\zeta_1 t) + c_2 \exp(\zeta_2 t) \end{bmatrix}$$

with the constants of integration, c_1 and c_2 determined by

$$c_1 = \frac{N^P(0)v_1^2 - N^R(0)}{v_1^2 - v_1^1}, \quad c_2 = \frac{N^R(0) - N^P(0)v_1^1}{v_1^2 - v_1^1}$$

where $N^R(0) > 0$ and $N^P(0) > 0$ are given. From this solution the dynamics of all other variables can be obtained.

2.A.6 Quadratic Preferences

This appendix states the systems of equations describing the full- and part-trade steady states under the assumption of quadratic preferences. We keep all assumptions about the economic environment and technology.

Full-trade Steady State

The following system of 6 equations in the 6 unknowns g, r, λ_R, λ_P, ω^R, m characterizes the full-trade steady state. Note that λ_i denotes the Lagrange multiplier of the optimization problem for a household in country i. We chose $w^P(t)b^P(t) = 1$ as the numeraire.

The Euler equation for both countries is determined by

$$g = r - \rho.$$

The resource constraints for the rich country, and the poor country, respectively, are given by

$$\begin{aligned} (1-\beta)L\theta_R &= g\,(1-m)\,F^R + (1-m)\,b^R(1-\beta)L\frac{1}{2}\left(s - \lambda_R\omega^R b^R\right) \\ &+ (1-m)\,\tau b^R \beta L \frac{1}{2}\left(s - \lambda_P \tau\omega^R b^R\right) \\ \beta L\theta_P &= gmF^P + mb^P\beta L\frac{1}{2}\left(s - \lambda_P\right) + m\tau b^P(1-\beta)L\frac{1}{2}\left(s - \lambda_R\tau\right). \end{aligned}$$

Free entry in both regions implies the following zero-profit conditions

$$\begin{aligned} \omega^R F^R &= \frac{1}{r}\left[\frac{\left(s - \lambda_R\omega^R b^R\right)^2(1-\beta)L}{4\lambda_R} + \frac{\left(s - \lambda_P\tau\omega^R b^R\right)^2\beta L}{4\lambda_P}\right] \\ \frac{F^P}{b^P} &= \frac{1}{r}\left[\frac{\left(s-\lambda_P\right)^2\beta L}{4\lambda_P} + \frac{\left(s - \lambda_R\tau\right)^2(1-\beta)L}{4\lambda_R}\right]. \end{aligned}$$

The balance of payments is determined by

$$\begin{aligned} 0 &= (1-m)\frac{1}{4}\left(\tau\omega^R b^R + \frac{s}{\lambda_P}\right)\beta\left(s-\lambda_P\tau\omega^R b^R\right) \\ &- m\frac{1}{4}\left(\tau+\frac{s}{\lambda_R}\right)(1-\beta)\left(s-\lambda_R\tau\right) \\ &- \beta T \end{aligned}$$

where $T \equiv T_P(t)/N(t)$.

Part-trade Steady State

The following system of 7 equations in the 7 unknowns g, r, m, q, ω^R, λ_P, and λ_R describe the part-trade steady state.

The Euler equation for the rich and poor country is given by

$$g = r-\rho.$$

Labor market clearing in the rich and poor country, respectively, implies

$$\begin{aligned} (1-\beta)L\theta_R &= g(1-m)F^R q(1-m)b^R L\left[(1-\beta)\left[s-\lambda_R\tau\frac{1}{2}\left(\tau\omega^R b^R+\frac{s}{\lambda_P}\right)\right]\right. \\ &+ \left.\tau\beta\frac{1}{2}\left(s-\lambda_P\tau\omega^R b^R\right)\right] \\ &+ (1-q)(1-m)b^R(1-\beta)L\frac{1}{2}\left(s-\lambda_R\omega^R b^R\right) \\ \beta L\theta_P &= gmF^P + mb^P\beta L\frac{1}{2}\left(s-\lambda_P\right) + m\tau b^P(1-\beta)L\left[s-\lambda_R\tau\frac{1}{2}\left(1+\frac{s}{\lambda_P}\right)\right]. \end{aligned}$$

The zero-profit conditions for firms in the rich and poor country, respectively, are given by

$$\begin{aligned} \omega^R b^R &= \frac{1}{r}\left[\frac{\left(s-\lambda_R\omega^R F^R\right)^2(1-\beta)L}{4\lambda_R}\right] \\ \frac{F^P}{b^P} &= \frac{1}{r}\left[\frac{\left(s-\lambda_P\right)^2\beta L}{4\lambda_P}+\frac{\tau\left(s-\lambda_P\right)\left[s-\lambda_R\tau\frac{1}{2}\left(1+\frac{s}{\lambda_P}\right)\right](1-\beta)L}{2\lambda_P}\right]. \end{aligned}$$

Arbitrage opportunities imply that firms located in the rich country that serve only the households in the rich country and firms serving all households in both countries must earn equal profits

$$\begin{aligned} \frac{\left(s-\lambda_R\omega^R b^R\right)^2(1-\beta)}{4\lambda_R} &= \left[\frac{\tau}{2}\left(\tau\omega^R b^R+\frac{s}{\lambda_P}\right)-\omega^R b^R\right](1-\beta)\left[s-\frac{\lambda_R\tau}{2}\left(\tau\omega^R b^R+\frac{s}{\lambda_P}\right)\right] \\ &+ \frac{\left(s-\lambda_P\tau\omega^R b^R\right)^2\beta}{4\lambda_P}. \end{aligned}$$

The balance of payments requires that

$$\begin{aligned} 0 &= q\,(1-m)\,\frac{1}{4}\left(\tau\omega^{R}b^{R}+\frac{s}{\lambda_P}\right)\beta\left(s-\lambda_P\tau\omega^{R}b^{R}\right) \\ &- m\tau\frac{1}{2}\left(1+\frac{s}{\lambda_P}\right)(1-\beta)\left[s-\lambda_R\tau\frac{1}{2}\left(1+\frac{s}{\lambda_P}\right)\right] \\ &- \beta T. \end{aligned}$$

3 A Dynamic North-South Model with Demand-Induced Product Cycles

Joint with Reto Foellmi and Sandra Hanslin

3.1 Introduction

In his seminal article Vernon (1966) explained international trade patterns with product cycles. He hypothesized that new goods would be introduced in countries with high per capita incomes (catering to the needs of such a market), after a while demand for these goods emerges abroad (as incomes grow) and they are exported. Later on, goods are imitated by less advanced countries, which have a relative cost advantage, such that the production moves there. Hence, goods that were once exported by rich countries are eventually imported by them. In a follow-up paper, Vernon (1979) explicitly emphasized the role of the demand side in shaping the typical product cycle:

> In the early part of the post-war period, the US economy was the repository of a storehouse of innovations not yet exploited abroad, innovations that responded to the labour-scarce high-income conditions of the US market. As the years went on, other countries eventually achieved the income levels and acquired the relative labour costs that had prevailed earlier in the United States. As these countries tracked the terrain already traversed by the US economy, they developed an increasing demand for the products that had previously been generated in response to US needs. That circumstance provided the consequences characteristically associated with the product cycle sequence . . . (Vernon 1979, p. 260).

It has been several decades since Vernon stressed the importance of the demand side for product cycles.[1] However, this is the first contribution to the literature, we are aware of, that provides a demand-based dynamic model which is able to generate the three stages of the product cycle described by Vernon (1966): (1) a product is exclusively produced and consumed in the North, (2) a product is produced in North and exported to South and (3) a product is imitated and exported from South to North. This chapter describes a dynamic general-equilibrium model of two regions, a wealthy North, and a poor South. Households have non-homothetic preferences over differentiated products such that consumption patterns differ across regions. In particular, households in North can afford to consume more and

[1] Two early studies of the product cycle that deserve mention can be found in Wells (1966) and Hirsch (1967), both doctoral students of Vernon at Harvard University.

newer products than households in South. Monopolistic firms in North innovate new products (horizontal innovations) whereas competitive firms in South randomly target Northern products for imitation. Trading products across regions is costless (the implications of this simplifying assumption are discussed in Section 3.4.4). In the steady state, products go on average through the following cycle. A new product is developed and introduced in North. Only after a certain time have households in South become rich enough to afford a "new" product that is produced in North. This demand lag is endogenously determined and depends, ceteris paribus, positively on the degree of inequality across regions and negatively on the innovation rate.[2] In other words, if Southern households, relative to Northern households, are poor the demand lag is long. Similarly, if incomes grow at a low rate the demand lag is long too. At this stage of the product cycle North is exporting the product. As time elapses further South eventually masters the technology to manufacture the product itself. Southern firms chose at random Northern products to imitate that have not yet been copied. They must invest resources in order to reverse engineer the production process of the chosen product. Once they have invested the necessary resources, they enter into price competition with the innovating firm in North. Because they have a cost advantage due to lower wages, they can underbid the Northern innovator and capture the whole market. Hence, South becomes an exporter of that product. The average time span a product is being manufactured in North is determined endogenously. In sum, we get on average a product cycle as described by Vernon (1966): At the stage of a new product, products are manufactured and consumed in North, at the mature stage they are exported to South, and eventually, at the stage of the standardized product they are manufactured in South and exported to North.

The remainder of the chapter is structured as follows. In Section 3.2, we distinguish our contribution from the existing literature. Section 3.3 presents suggestive evidence for the product cycle described in the introductory Section 3.1 by studying six major consumer durables. In Section 3.4, we introduce the model and solve for the steady state, and transitional dynamics. Comparative statics results of changes in Southern productivity, and changes in inequality across regions are discussed in Section 3.5. Section 3.6 extends the model towards hierarchic preferences, and learning-by-doing. Eventually, Section 3.7 concludes.

3.2 Related Literature

The theory of Vernon (1966) grew out of dissatisfaction with classical trade theories, which explain trade between countries with differences in relative factor endowments (Heckscher-Ohlin) or differences in relative productivities (Ricardo). On the one hand, these theories missed characteristics like countries' per capita incomes (Burenstam-Linder 1961) that are thought to be important determinants of international trade (see e.g. Markusen 1986; Fieler

[2] Our use of the term "demand lag" differs from Posner (1961). He thinks of the demand lag as the delay in the acceptance of foreign goods in the domestic market, i.e. foreign goods might not be considered perfect substitutes for home-produced goods until some time elapses. We define the demand lag as the time it takes in the poor South for incomes to grow sufficiently such that households there can afford to buy goods produced in North, abstracting from differences in tastes.

2011), and on the other hand, struggled to explain observed trade flows (Leontief 1953, 1956). Since Vernon put forward his verbal theory of the product cycle, a number of economists have both formalized the product cycle theory in theoretical models as well as put it to empirical tests.

One of the first to study product cycles in a theoretical model was Krugman (1979). In his model, an advanced North introduces new products at a constant exogenous rate, i.e. the product space expands, and a less advanced South copies these goods also at a constant exogenous rate. Higher per capita income in North depends on quasi rents from the Northern monopoly in new goods, i.e. North must continually innovate to maintain its relative and absolute position. Later, Grossman and Helpman (1991a) extended Krugman's (1979) model, and endogenized innovation and imitation rates. In their model, long-run growth is faster the larger the resource base of the South and the more productive its resources in learning the production process. The reason is that profits during the monopoly phase are higher when a smaller number of Northern firms compete for resources in the manufacturing sector, which outweighs the effect of a higher risk-adjusted interest rate since profits accrue on average for a shorter period of time. Both models focus on the supply-side aspect of the product cycle theory, i.e. how the diffusion of technology and the determination of relative wages depend on technology and preference parameters of the model (for a more recent example see Acemoglu et al. 2012). However, in both models, demand patterns in North and South are identical because agents have homothetic constant-elasticity-of-substitution (CES) preferences. In other words, the consumption basket demanded in North is simply an inflated clone of the one in South. "This is clearly at odds with the fact, stressed by Vernon, that new goods are not typically consumed in the South until later in the cycle" (Stokey 1991, p. 63). Hence, Stokey (1991) focuses on the demand side. In her static model with vertically differentiated goods North manufactures high-quality products whereas South manufactures low-quality products. Since agents have non-homothetic preferences all of the products manufactured in the North are consumed domestically but only the lower-quality products are exported. Stokey (1991) is interested in the effect of population size and productivity changes on relative wages, production and trade patterns, respectively, and the terms of trade. Kugler and Zweimüller (2005) build a dynamic North-South model where households have non-homothetic preferences. Their model is close to our setup. However, the model in Kugler and Zweimüller (2005) is not a full-fledged general-equilibrium model since interest rates are exogenously determined. Furthermore, the focus of their analysis is on the cross-sectional composition of aggregate demand rather than on product cycles.

Our model differs from the existing literature in the following ways. In contrast to e.g. Stokey (1991) or Flam and Helpman (1987) we focus on horizontal instead of vertical innovations. In addition, we differ from Stokey (1991), Matsuyama (2000) or Falkinger (1990) who build static Ricardian trade models where agents have (hierarchic) non-homothetic preferences by developing a dynamic general-equilibrium model with monopolistic competition. We model the demand side as Foellmi and Zweimüller (2006), whereas the supply-side is borrowed from Grossman and Helpman (1991a). Incorporating non-homothetic preferences into these types of models enables us to formalize the product cycle hypothesis and analyze the effects of the demand side on the product cycle.

3.3 Motivation

3.3.1 Empirical and Anecdotal Evidence

There have been many attempts to test the product cycle hypothesis empirically. Among the first who found evidence for the product cycle theory were Wells (1969) for consumer durables and Hirsch (1967) for electronic products. Hirsch (1975) and Mullor-Sebastian (1983) find that industrial product groups behave according to the product cycle theory. More recently, Feenstra and Rose (2000) find evidence for product cycles by ranking goods (and countries) according to the year they are first exported (export) to the United States. In particular, they show that less sophisticated goods like furniture are imported early by the US, and that advanced countries like Canada, UK, Germany, Japan and France start to export early to the US. Furthermore, there is a negative correlation between real GDP per capita and the country ranking, i.e. high-income countries tend to have a low ranking, which means that they start exporting early to the US (i.e. are more advanced).

Besides empirical evidence there is anecdotal evidence for the product cycle hypothesis, e.g. products like color televisions, computer games, or electric can openers seem to follow or have followed Vernon's product cycle. A typical example that is currently in an early stage of the life-cycle is the Tesla Roadster (the first all electrical sports car) which is produced by Tesla Motors Inc. in California, USA. It was exclusively available in the US for around two years before it became available in some Western European countries, Hong Kong, Japan and Australia. The product cycle theory predicts that demand abroad further increases as incomes grow, and eventually the car is being imitated and production is moving abroad where manufacturing costs are comparatively lower. In the following section, we are going to illustrate the product life-cycle with 6 major consumer durables of the 20th century such as the microwave oven, dishwasher, freezer, washing machine, dryer and VCR.

3.3.2 The Product Cycle of 6 Major Consumer Durables

Instead of attempting to empirically test the product cycle hypothesis, we take a different route and provide suggestive evidence by looking at three distinct features characterizing the product cycle described by Vernon (1966) that should be observed in the data.[3] First, new products are not introduced in all countries simultaneously, with the lag of introduction varying negatively with GDP per capita. Second, as the production of goods migrates from North to South, North should start out as a net exporter of a given product, and over time become a net importer of that good. Third, production of a given good should be high in developed relative to developing countries early in the product cycle, and low later on. We interpret the patterns

[3] A conclusive analysis would require time series data on production and consumption at the product level across a large sample of countries - the gathering and analysis of such data is beyond the scope of this chapter.

found in the data for these 6 consumer durables as suggestive evidence for the type of product cycles Vernon (1966) had in mind. In particular, we believe they are representative for the type of final consumer goods that the product cycle hypothesis is relevant. It is obvious that the theory does not apply to trade in e.g. agricultural commodities or natural resources.

We study 6 major consumer durables introduced in the 20th century, i.e. dishwasher, dryer, freezer, microwave oven, VCR, and washing machine. For each consumer durable we know the year of introduction in 16 European countries, i.e. Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and United Kingdom. The dates of introduction for these consumer durables were kindly provided by Tellis et al. (2003).[4] Trade data are U.S. import and export data at the 5 digit SITC level from 1972-2006, which are provided by the Center for International Data at UC Davis (Feenstra 1996, 1997). GDP data are taken from Penn World Tables (PWT) 7.0, see Heston et al. (2012). Production data are obtained from the Industrial Commodity Production Statistics Database (United Nations Statistics Division 2012). Internet research reveals that all of these products were first introduced in the United States: the electric dishwasher was introduced in 1950 (Hobart Corp.), the automatic electric clothes dryer in 1949 (Hamilton Manufacturing Corp., General Electric), the domestic deep freezer in 1949 (General Electric), the countertop microwave oven in 1967 (Amana Corp.), the VCR in 1965 (Sony, Ampex, RCA), and the automatic electric washing machine in 1947 (Bendix, General Electric).

Demand Lags

Table 3.1 shows the year of introduction, defined by Tellis et al. (2003) as the first year commercial sales for the corresponding product were registered, of the 6 consumer durables, and GDP per capita (PWT 7.0, PPP, 2005 USD) in the year the product was introduced in the United States. For example, the countertop microwave oven was first introduced in the US in 1967, and last introduced in Greece and Portugal in 1982. GDP per capita in 1967 was USD 19,522 in the US, whereas in 1967 Greece had a GDP per capita of USD 9,742 and Portugal one of USD 5,937.[5] We observe that the year of introduction varies across countries. It appears that on average, products were first introduced in countries with a high GDP per capita like the US and UK, and last introduced in countries with a low GDP per capita like Greece, Portugal and Spain. This is also suggested by Spearman's rank correlation coefficient between the year of introduction and GDP per capita in that year, shown at the bottom of Table 3.1.

Let us look closer at the microwave oven, which we consider a typical household appliance. In 1946, Percy Spencer, an American engineer, while working on radar technology for the U.S. defense company Raytheon Corporation accidentally discovered that microwaves are capable of heating food extremely quickly. The story goes that a candy bar in Spencer's pocket melted during an experiment. Spencer realized the commercial potential, especially for a high-income

[4]Unfortunately, we don't have data on the diffusion of the 6 consumer durables.

[5]Note that PWT 7.0 provides data from 1950-2009. Hence, we use data on GDP per capita in 1950 to approximate GDP per capita in the years 1947 and 1949.

Country	Dishwasher		Dryer		Freezer		Microwave Oven		VCR		Washing Machine	
	Year	GDP pc	Year	GDP pc	Year	GDP pc	Year	GDP pc	Year	GDP pc	Year	GDP pc
Austria	1962	6296	1965	6296	1953	6296	1974	13575	1977	12651	1962	6296
Belgium	1960	7992	1966	7992	1956	7992	1974	13602	1975	12820	1955	7992
Denmark	1960	9366	1965	9366	1954	9366	1974	15980	1977	15145	1958	9366
Finland	1964	6192	1973	6192	1961	6192	1975	11590	1978	11203	1960	6192
France	1959	7107	1963	7107	1960	7107	1975	13578	1974	12522	1954	7107
Germany	1960	6251	1966	6251	1956	6251	1969	14348	1974	14377	1952	6251
Greece	1966	2576	1990	2576	1972	2576	1982	9742	na	na	1964	2576
Ireland	1965	5880	1963	5880	1958	5880	1976	8789	na	na	1966	5880
Italy	1961	5361	1968	5361	1965	5361	1975	12305	1976	11015	1957	5361
Netherlands	1960	9961	1968	9961	1960	9961	1971	16356	na	na	1960	9961
Norway	1961	9434	1970	9434	1957	9434	1976	16366	1977	14966	1960	9434
Portugal	1966	2614	1973	2614	1956	2614	1982	5937	na	na	1956	2614
Spain	na	na	1973	3796	1972	3796	1973	10215	1977	9321	1964	3796
Sweden	1959	10301	1969	10301	1953	10301	1973	17043	1977	16380	1958	10301
Switzerland	na	na	1966	13712	na	na	1973	22880	1978	22056	na	na
United Kingdom	1958	10447	1960	10447	1954	10447	1971	14886	1974	14365	1954	10447
United States	1950	13119	1949	13119	1949	13119	1967	19522	1965	18364	1947	13119
Spearman's corr coeff	-0.858		-0.545		-0.675		-0.651		-0.011		-0.463	

Source: Tellis et al. (2003), PWT 7.0

Table 3.1: Introduction of 6 major consumer durables across European countries

market like the US, of his discovery, and Raytheon filed for patents. In 1947, Raytheon produced the first commercial microwave oven named "Radarange", which was sold to restaurants etc. Twenty years later, in 1967, Amana, a division of Raytheon, introduced the first domestic countertop microwave oven, marking the beginning of the use of microwave ovens in American kitchens (see e.g. Osepchuk 1984).

Figure 3.1 below shows the relationship between the demand lag in years of the countertop microwave oven relative to the US across the 16 European countries, and their GDP per capita relative to the United States in 1967 (PPP, 2005), the year the microwave oven was introduced in the US. We observe that on average the lower a countries' GDP per capita relative to the US in 1967, the longer the time span until households in a country start purchasing the microwave oven. For example, the Netherlands had a GDP per capita in 1967 that was about 20 percent lower than the US and households started consuming microwave ovens 4 years later than the US. By comparison, Portugal's GDP per capita in 1967 was about 30 percent of that in the US and households began purchasing microwave ovens 15 years later. Note that countries below the line of fit have a higher average growth rate between 1950-1990 than countries above the line of fit (calculations based on PWT 7.0). Switzerland is an extreme outlier in the sense that even though its GDP per capita in 1967 was higher than the US GDP per capita, households first purchased the microwave oven 6 years later than households in the US.

The graphs for the other consumer durables look similar. For each country Figure 3.1 also plots the average of GDP per capita relative to the US in the year of introduction across all 6 consumer durables against the average lag in years in the introduction of these consumer durables. We conclude from Figure 3.1 that there is a negative correlation, suggesting that on average, in countries where households enjoy a high income, these consumer durables are

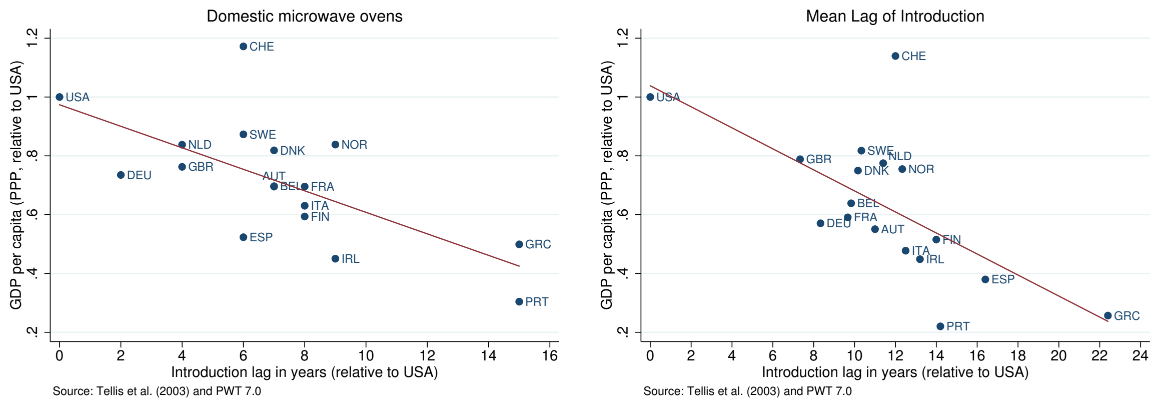


Figure 3.1: Relationship between GDP per capita and demand lag for the microwave oven (left-hand side) and across all 6 consumer durables (right-hand side)

purchased sooner.[6] Again, we notice that Switzerland is an extreme outlier that might be explained by its lacking integration into the European market at that time, its relatively highly regulated domestic market or its small population size. Hence, we also take into account (relative) population size, and for each product category j as well as across all product categories, estimate the following model by OLS:

$$\log\left(\Delta_{ij}\right) = \beta_0 + \beta_1 \log\left(\frac{GDPpc_{ij}}{GDPpc_{US}}\right) + \beta_2 \log\left(\frac{Pop_{ij}}{Pop_{US}}\right) + \varepsilon_{ij}.$$

In other words, we regress the introduction lag Δ (in logs) in each of the 6 product categories j on the (log of) GDP per capita in country i relative to the GDP per capita of the US and the (log of) population size of country i relative to the population size of the US, both in the year the US introduced the product category. The coefficient β_1 shows the importance of (relative) GDP per capita, holding relative population sizes constant. We expect β_1 to be negative. Table 3.2 below illustrates that (relative) GDP per capita has a negative effect on the introduction lag of all 6 products, controlling for (relative) population sizes.

Export Performance

We now turn to the export performance of the United States in these product categories. In particular, we look at the value of US exports (X_{jt}) in product category j at time t relative to the value of US exports plus imports (I_{jt}) in category j at time t, i.e. $\widetilde{x}_{jt} \equiv X_{jt}/(X_{jt} + I_{jt})$. We note that if $\widetilde{x}_{jt} \in (0, 0.5)$ the US is a net importer in product category j at time t, and if $\widetilde{x}_{jt} \in (0.5, 1)$ the US is a net exporter. The product cycle hypothesis offers an explanation for a falling export ratio $\widetilde{x}_{jt}$ in the data. The US should start out as a net exporter and become a net importer over time in a given product category. Again, in Figure 3.2, we take a look at the

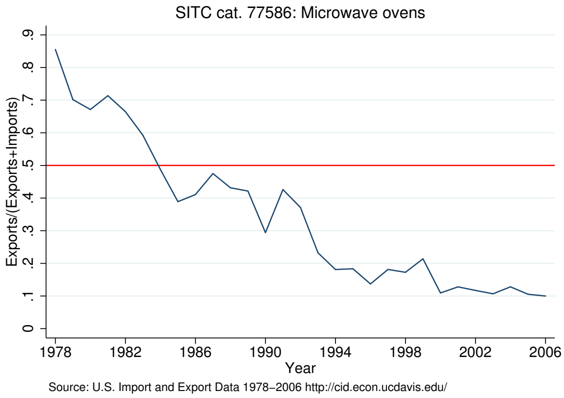
[6] We expected GDP per capita to be more important for product take-off (i.e. when a certain threshold of sales has been reached) than for the time of introduction. However, it seems that firms base their decisions to launch new products (and form their expectations about future sales performance) as much on the general level of development in a country (e.g. high average income level) as on the existence of a small group of rich people.

Table 3.2: Correlation between (log) relative GDP per capita and (log) introduction lag Δ

	log($\Delta mean$)	log($\Delta dish$)	log($\Delta dryer$)	log($\Delta freeze$)	log($\Delta micro$)	log(Δvcr)	log($\Delta wash$)
log(rel GDPpc)	-0.428	-0.399	-0.427	-0.702	-0.848	-0.124	-0.249
	(-3.95)	(-9.75)	(-3.61)	(-2.49)	(-2.88)	(-0.88)	(-1.45)
log(rel pop)	-0.109	-0.107	-0.098	0.094	-0.221	-0.108	-0.235
	(-2.41)	(-6.03)	(-1.77)	(0.75)	(-2.48)	(-3.86)	(-3.09)
adj. R^2	0.546	0.911	0.460	0.262	0.460	0.547	0.399
#obs	16	14	16	15	16	12	15

Notes: t-values in parentheses

export performance of the United States in the product category of microwave ovens, as well as across all 6 consumer durables, both across the 16 European countries.



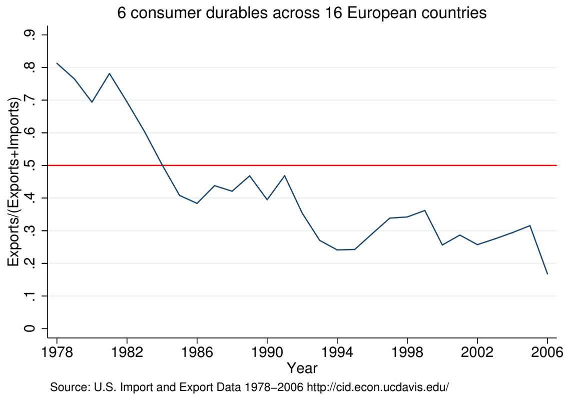


Figure 3.2: US export ratio in microwave ovens across 16 European countries (left-hand side) and across all 6 consumer durables (right-hand side)

We observe that the US starts out as a net exporter of microwave ovens at the beginning of the sample period in 1978 and ends up as a net importer at the end of 2006, switching around 1983/84. A possible interpretation for the decline in the export ratio is that firms in the 16 European countries mastered the technology to produce microwave ovens, and due to lower production costs were able to compete with US firms in their home markets as well as in the US market. In other words, US firms became less competitive in their export markets and/or European firms became more competitive in the US market, such that US exports relative to U.S. imports decreased. The export performance in the other products is similar, except for the domestic deep freezer where the export ratio $\widetilde{x}_{jt}$ follows an inverse U-shape over 1978-2006. The right-hand side panel in Figure 3.2 shows the export performance of the U.S. in the 16 European countries aggregated over the 6 product categories. We see that the value of US exports relative to its imports also declines across all 6 consumer durables combined.

Production Patterns

Eventually, for the microwave oven and the washing machine we study production data for the United Kingdom, the United States, Japan, South Korea, Brazil, Russia, India, China, and Argentina during the time period of 1982-2008. The United Nations Commodity Statistics Yearbook (United Nations Statistics Division 2012) collects data on production of industrial commodities by country.[7] Unfortunately, we don't have historical data on the production of the microwave oven in China and India. Figure 3.3 plots the number of units (in millions) relative to US production for the microwave oven (left-hand side) and the washing machine (right-hand side) over the time period of 1982 to 2008.

We make the following observations. First, U.S. production of both consumer durables is declining from the 1980s until 2008 relative to emerging countries like e.g. Brazil, Russia, India and China. Second, while emerging countries are catching up, the production of developed countries like the UK and Japan only moderately increases or even decreases. The relative increase in production of the washing machine is especially strong for China. However, also Indian production of washing machines is catching up slowly. Again, this is consistent with the product cycle hypothesis, which suggests that the production of microwave ovens and the washing machine should move from developed countries to developing countries as firms in these countries acquire the technology to produce microwave ovens and have lower production costs. Data limitations prevent us from investigating the production patterns for the rest of the consumer durables discussed above.

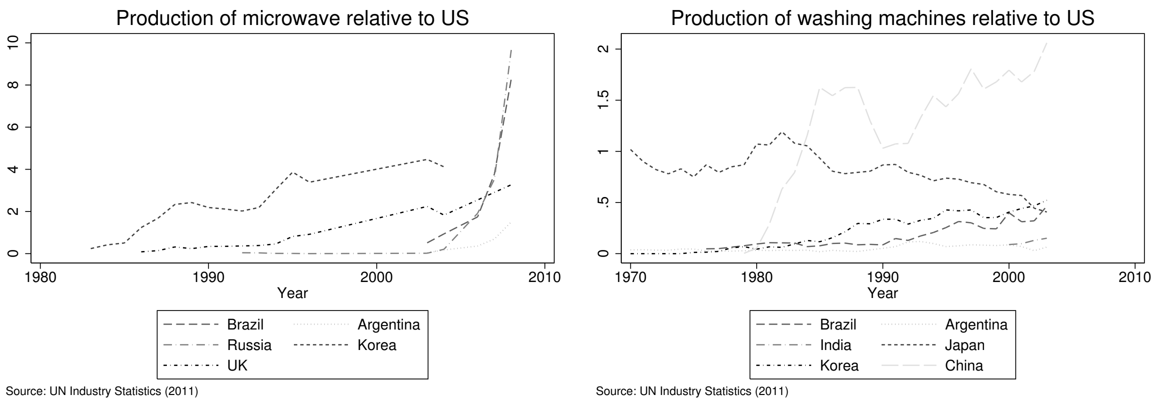


Figure 3.3: Production of microwave oven and washing machine relative to US

[7]The data is collected through annual questionnaires sent to national statistical authorities. The data reported by the United Nations Commodity Statistics Yearbook reflect volume (and value) of production sold during the survey period, which is defined as the production carried out at some time, which has been sold (invoiced) during the reference period.

3.4 Model

3.4.1 Distribution and Endowments

The economy consists of two regions $i \in \{N, S\}$, an industrialized North (N) and a less developed South (S). The population size of the economy is L, a fraction β lives in the South and a fraction $(1-\beta)$ in the North. We assume that each household regardless of its residence inelastically supplies one unit of labor on the local labor market. This implies that aggregate labor supply in the South is given by βL, and by $(1-\beta)L$ in the North. Furthermore, suppose that each household holds domestic and foreign assets. Hence, income inequality is endogenously determined and originates from differences in labor and capital incomes across countries.

In order to study ceteris paribus effects of income inequality across countries we introduce a transfer system (e.g. foreign aid) between North and South so that each household in the North pays/receives a lump-sum tax/benefit $T_N(t)$, respectively $T_S(t)$. The transfer system must run a balanced budget in each period so that $(1-\beta)LT_N(t) = \beta LT_S(t)$, and transfers grow at the same rate as incomes. We will take $T_S(t)$ as the exogenous variable so that through the balanced budget condition $T_N(t)$ is endogenously determined.

3.4.2 Preferences

There is a continuum of differentiated products in the economy indexed by $j \in [0, \infty)$, where only a subset $N(t)$ is available on the market at each point in time. We assume differentiated products to be indivisible, and model consumption as a binary decision. Hence, households consume either 1 unit of product j at time t, or they don't consume that product at all. Instantaneous utility is non-homothetic and takes the following form

$$u\left(\{c(j,t)\}_{j=0}^{N(t)}\right) = \int_0^{N(t)} c(j,t)\, dj \tag{3.1}$$

where $c(j,t)$ is an indicator function that takes the value one if product j is consumed at time t, and zero otherwise. The indicator function $c(j,t)$ will be specific to the income group, i.e. the region. The specification of the instantaneous utility function contrasts with the constant-elasticity-of-substitution (CES) form as follows. With zero-one preferences households can only choose consumption along the extensive margin, i.e. choose how many different products they want to purchase, whereas with CES preferences they can only choose consumption along the intensive margin, i.e. how many units of each product they want to buy. In that sense, our preferences are no less special or general than CES preferences.[8] Furthermore, note that

[8]For the sake of illustration, suppose that the whole product set available to households consists of the six consumer durables in Section 3.3. With the preferences specified in (3.1) wealthy households in the North would consume one unit of all consumer durables available whereas poor households in the South could not afford to consume all goods available, and for example could only purchase one washing machine and one freezer (some of the "older" goods available). With CES preferences Northern and Southern households would both consume all six consumer durables available. However, Northern households would purchase e.g. five units of each good whereas Southern households could only buy one unit each.

preferences in (3.1) are symmetric, i.e. no product is intrinsically better or worse than any other product. In other words, there is no explicit consumption hierarchy. This allows us to order products in ascending order from old to new, such that product j is developed before product j', where $j' > j$.[9]

The household's intertemporal objective function is given by

$$U(0) = \int_0^\infty \exp(-\rho t) \log u \left(\{c(j,t)\}_{j=0}^{N(t)} \right) dt \tag{3.2}$$

where $\rho > 0$ denotes the time preference rate. Note that intertemporal preferences given by (3.2) are homothetic. Households maximize their lifetime utility (3.2) subject to non-negativity constraints $c(j,t) \geq 0$ for all j, t, and to their lifetime budget constraint

$$\int_0^\infty \int_0^{N(t)} p(j,t)c(j,t)dj \exp\left(-\int_0^t r(s)ds\right) dt \leq a(0) + \int_0^\infty (w(t) + T(t)) \exp\left(-\int_0^t r(s)ds\right) dt$$

where $r(t)$ denotes the risk-free interest rate, $a(0)$ initial wealth, and $w(t)$ the market clearing wage rate. The solution to the household problem has been relegated to Appendix 3.A.1. From the maximum principle conditions we derive the individual Marshallian demand function for product j:

$$c(j,t) = \begin{cases} 1 & p(j,t) \leq z(j,t) \\ 0 & p(j,t) > z(j,t) \end{cases} \tag{3.3}$$

where $z(j,t) \equiv [u(\cdot)\lambda(t)]^{-1}$ denotes the willingness to pay. Figure 3.4 below shows the individual demand curve (3.3) for product j. The costate variable, which can be interpreted as marginal utility of wealth at time t, is denoted by $\lambda(t)$. Households purchase one unit of a product if the price of that product does not exceed their willingness to pay. Since preferences are symmetric over all products the willingness to pay is identical for all products j. However, the willingness to pay depends on $\lambda(t)$, i.e. on the shadow price of (lifetime) income. Hence, consumption patterns differ across regions since by our distributional assumptions (lifetime) incomes are different across regions. Wealthy households in the North, with a lower equilibrium value of $\lambda(t)$, consume a larger set of products than poor households in the South.

[9] Note that the same ordering would emerge if we assumed instantaneous utility to take the following form $u(c(j,t)) = \int j^{-\eta} c(j,t)\, dj$. The power function $j^{-\eta}$ implies that (instantaneous) marginal utility is falling in the index j, i.e. higher indexed goods yield lower marginal utility than lower indexed goods. The parameter $\eta \in (0,1)$ determines the "steepness" of the hierarchy, i.e. how fast marginal utility falls in index j. With these preferences households start consuming low-indexed goods (as they yield higher marginal utility) and expand consumption towards high-indexed goods until their income is used up. To keep the model simple, we will assume that such a hierarchy in consumption latently exists rather than explicitly modeling it. For a detailed discussion see Section 3.6.1.

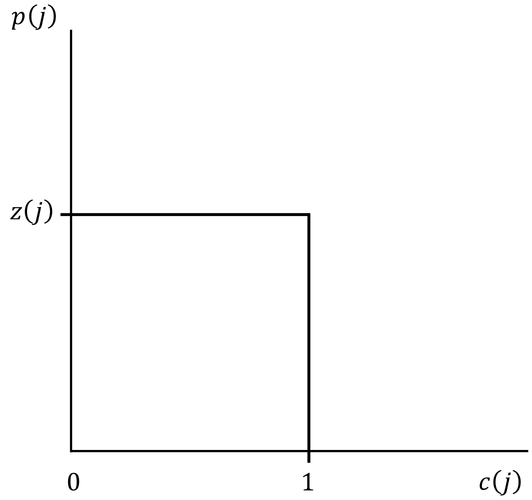


Figure 3.4: Individual demand

3.4.3 Technology and Trade Integration

Innovation Technology in North

New products are designed and developed in high-income countries.[10] Each firm in the North is a single-product firm, which has access to the same innovation technology. The creation of a new product requires $F^N(t) = F^N/N(t)$ units of labor, once this set-up cost has been incurred, the firm has access to a linear technology that requires $b^N(t) = b^N/N(t)$ units of labor to produce one unit of output, with $F^N, b^N > 0$ being positive constants. Innovations obey an important spillover because they imply technical progress. We assume that the knowledge stock of this economy equals the number of known designs $N(t)$. The labor coefficients are inversely related to the stock of knowledge. New products are protected by infinite patents but face a positive probability of being copied by a Southern firm (patent infringement). We assume that firms in the North cannot license technology to Southern firms, or set up manufacturing plants in the South (i.e. engage in foreign direct investment).

Imitation Technology in South and Transportation Costs

As in Grossman and Helpman (1991a) we assume that each new product, which has been developed in the North at time t, faces the same positive probability of being imitated by a Southern firm at some time $\widetilde{T} > t$. At the time the product is developed date $\widetilde{T}$ is unknown. In other words, $\widetilde{T}$ is a random variable that represents the age of a product at the time of imitation. A Southern firm selects at random one of the existing products in the North, which

[10] In principle, one could think that both North and South have access to the innovation technology but that the South is sufficiently unproductive at developing new products compared to the North such that in equilibrium no innovation takes place in the South. Since it is difficult to measure research productivity, for illustration's sake, consider research input. World Bank (2012) data on research and development spending of low/middle and high income countries show that high income countries on average spent about 2.5 times as much on R&D in percent of their GDP than low and middle income countries during the period 2000-2007.

has not yet been copied, for imitation. We assume that firms in the South benefit in reverse engineering and production from the total stock of knowledge (i.e. there are international knowledge spillovers). Imitation of a selected product requires $F^S(t) = F^S/N(t)$ units of labor, with $F^S > 0$. Investing $F^S(t)$ allows a Southern firm to learn the production process of the randomly chosen product with probability one. Hence, there is complete certainty for a Southern imitator that reverse engineering succeeds. Subsequent production of the copied good requires $b^S(t) = b^S/N(t)$ units of labor per unit. Finally, we assume that product markets are fully integrated and trade costs are zero.

3.4.4 Equilibrium

Depending on parameter values, two decentralized equilibria can emerge: (i) households in the South are too poor to afford any Northern products or (ii) they can afford at least some Northern products. In case (i) no trade equilibrium exists. Hence, we focus on the interesting case (ii), and assume in the following that households in the South can afford some Northern products. In proving the existence of the equilibrium, we will derive the necessary assumption on parameters. Let us denote the set of all products available in the economy as $N(t) = N^N(t) + N^S(t)$, where $N^N(t)$ denotes the subset of products that have not yet been imitated by the South, and $N^S(t)$ the subset of products that have been copied by the South.

World Demand

In the equilibrium we consider households in the North consume all products available in the market $N_N(t) = N(t)$, whereas households in the South consume only a subset of all products $N_S(t) \subset N(t)$, which includes all products manufactured in the South and some but not all Northern products. World demand for product j can be derived by horizontally aggregating individual demand (3.3) across regions. It is determined by:

$$C(j,t) = \begin{cases} 0, & p(j,t) > z_N(j,t) \\ (1-\beta)L, & z_S(j,t) < p(j,t) \leq z_N(j,t) \\ L, & p(j,t) \leq z_S(j,t) \end{cases} \tag{3.4}$$

where $z_i(j,t)$, with $i \in \{N,S\}$, denotes the willingness to pay of households in the North, and South, respectively. Since the willingness to pay is the same for all products j, aggregate demand is the same for all products. World demand (3.4) is depicted in Figure 3.5 below.

If the price of a product exceeds the willingness to pay of Northern households, there is no demand for that product. With a price between the willingness to pay of Southern and Northern households only the latter purchase the product. If the price falls short of the willingness to pay of households in the South everyone purchases it. Figure 3.5 is drawn under the assumption that the willingness to pay of Southern households exceeds marginal costs $b^N(t)w^N(t)$, which holds true in the equilibrium of interest.

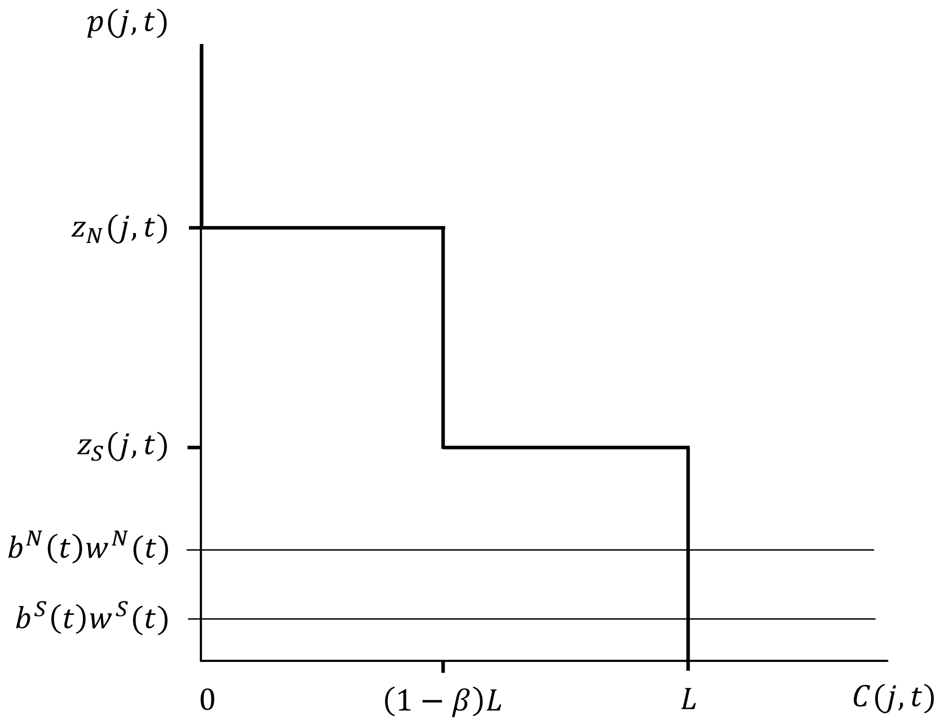


Figure 3.5: World demand

Aggregate Supply

Let us first consider the problem of a monopolistic firm j located in the North. Firm j maximizes operating profits

$$\pi^N(j,t) = \left[p(j,t) - w^N(t)b^N(t)\right] C(j,t) \tag{3.5}$$

subject to aggregate demand (3.4) by choosing a price $p(j,t)$ such that marginal revenue equals marginal cost. From Figure 3.5 and the discussion in the previous section it follows that there are two candidates for the price that maximizes profits (3.5). Firm j either sets a high price equal to the willingness to pay of Northern households $z_N(j,t)$ and sells exclusively to domestic households, or it sets a low price equal to the willingness to pay of Southern households $z_S(j,t)$ and serves both markets.

We assume that firms cannot price discriminate across regions. As there are no trade costs, arbitrageurs would take advantage of any price differential between North and South.[11] Thus, exporters set the same price in both regions. This implies that in equilibrium not all Northern firms export. To see this, suppose that at every point in time all Northern firms would set prices equal to the willingness to pay of Southern households and sell to everyone. In that case, households in the North would not exhaust their budgets, i.e. the shadow price of their (lifetime) income would become zero. That would imply an infinitely large willingness to pay

[11] The threat of arbitrage opportunities imposes a price setting restriction on firms. If there are no trade costs the price setting restriction is always binding. However, in the presence of iceberg trade costs the price setting restriction might not be binding. In particular, if the difference in per capita incomes between North and South were sufficiently low, all newly invented products in the North would be exported to the South right away.

for an additional product. Consequently, Northern firms had an incentive to deviate from selling to everyone and sell exclusively in the North. Hence, a situation where all Northern firms serve all households cannot be an equilibrium. Also, by the same argument, a situation where all Northern firms sell exclusively to Northern households cannot be an equilibrium as the willingness to pay of Southern households for a Northern product would become infinitely large.

In an equilibrium, where some Northern firms serve all households in both regions and others serve exclusively the domestic region, firms must be indifferent between selling only to Northern households and selling to all households at any point in time. Hence, the following arbitrage condition must hold

$$\left[z_N(j,t) - w^N(t)b^N(t)\right](1-\beta)L = \left[z_S(j,t) - w^N(t)b^N(t)\right]L. \tag{3.6}$$

In the aggregate, a measure n of firms sells in the North and South whereas $(1-n)$ firms sell only in the North. Due to symmetric preferences, however, the behavior of a single firm is indeterminate. Because we are free to order the different goods, we may think of the following firm behavior at the micro level that generates the described outcome at the macro level: After developing a new product each firm starts marketing its product solely in the North and after a certain period of time has elapsed, i.e. the time it takes for incomes in the South to have grown sufficiently, begins exporting. In that case, there are at any point in time new products that are sold exclusively in the domestic market and older products that are exported as well. Section 3.6 discusses two possible extensions where the product cycle at the firm level is determinate. We argue that while the model would become substantially more complex the basic structure and intuition of the baseline model is preserved.

The Northern firm, which develops product j at time t, faces a positive probability that its product will be copied by a Southern firm. After a product has been imitated, the Southern firm maximizes operating profits

$$\pi^S(j,t) = \left[p(j,t) - w^S(t)b^S(t)\right]C(j,t)$$

where $C(j,t) = L$ is given by (3.4). After the firm in the South has copied the Northern product j it enters into a price competition with the Northern firm currently producing j (the innovating firm). This forces the Southern firm to set a limit price equal to the marginal costs of the competing firm in the North. Hence, optimal prices of Southern products are equal to $w^N(t)b^N(t)$.[12]

[12] The wide-gap case discussed in Grossman and Helpman (1991a) where Southern firms can set the monopoly price cannot occur here since $z_S(t) > w^N(t)b^N(t) > w^S(t)b^S(t)$ in equilibrium as otherwise, no firm in the North would export to the South. Our case is similar to their narrow-gap case where Southern firms charge prices marginally below the marginal cost of Northern firms.

Labor Markets

Labor is immobile across regions but regional labor markets are assumed to be perfect. In particular, in the North labor is completely mobile between production and R&D, and in the South between production and reverse engineering. Labor market clearing in the North demands that

$$(1-\beta)\,L = g(t)F^N + b^N L\,[n(t)-m(t)] + (1-\beta)\,b^N L\,[1-n(t)] \tag{3.7}$$

where we defined $g(t) \equiv \dot{N}(t)/N(t)$, and the share of goods consumed and produced in the South, respectively, as $n(t) \equiv N_S(t)/N(t)$ and $m(t) \equiv N^S(t)/N(t)$. The first term in (3.7) on the right-hand side denotes labor demand from the R&D sector, the second term labor demand from the production of older Northern products consumed by all households in both regions, and the third term labor demand from the production of newer Northern products exclusively consumed by Northern households.

Similarly, labor market clearing in the South requires

$$\beta L = g^S(t)m(t)F^S + m(t)b^S L \tag{3.8}$$

where we defined $g^S(t) \equiv \dot{N}^S(t)/N^S(t)$. The right-hand side in (3.8) denotes labor demand from reverse engineering, and production of imitated products which are consumed by all households in both regions.

Capital Markets

We assume that international capital markets are perfect, hence, interest rates equalize across regions. The expected present discounted value of profits of product j that was introduced at time t is determined by equation (3.9) below, given the instantaneous rate of imitation $\mu(t) \equiv \dot{N}^S(t)/N^N(t)$. We make the standard assumption of free entry into product development in the North. Hence, the expected value of product j must equal R&D costs $w^N(t)F(t)$,

$$v^N\,(j,t) = \int_t^{\infty} \exp\left(-\int_t^s \left(r(\tau)+\mu(\tau)\right)d\tau\right)\pi^N\,(j,s)\,ds = w^N(t)F^N(t). \tag{3.9}$$

Note that profits are discounted using the risk-adjusted interest rate $r(\tau)+\mu(\tau)$, where $r(\tau)$ is the risk-free interest rate and $\mu(\tau)$ the risk premium. Since we assume capital markets to be perfect, households can diversify away the idiosyncratic risk of a Northern firm of being copied by holding a portfolio of shares in all Northern firms.

Free entry also prevails in the reverse engineering sector in the South, which is not an uncertain activity, so that their present discounted value of profits $v^S(j,t)$ must equal the imitation cost $w^S(t)F^S(t)$,

$$v^S\,(j,t) = \int_t^{\infty} \exp\left(-\int_t^s r(\tau)d\tau\right)\pi^S\,(j,s)\,ds = w^S(t)F^S(t). \tag{3.10}$$

Asset Holdings and Balance of Payments

The balance of payments in present value terms is determined by

$$\begin{aligned} 0 &= \int_0^\infty \left\{ \left[(1-\beta)LN^S(t)\omega^N b^N - \beta L \left(N_S(t) - N^S(t) \right) z_S(t) \right] \right. \\ &+ \left. \beta L T_S(t) \right\} \exp\left(-\int_0^t r(s)ds \right) dt \end{aligned} \tag{3.11}$$

where the first term in brackets on the right-hand side denotes the trade balance and the second term net transfer payments. We assume that net foreign assets (portfolio investments) are zero.[13] Note that if $T_S(t) > 0$ for all t, the South runs a (permanent) trade deficit, i.e. the value of its exports falls short of the value of its imports.

3.4.5 Steady State

The economy is in a steady state if Northern firms introduce new products at a constant rate g and Southern firms imitate at a constant rate μ. In steady state, shares of resources devoted to R&D and production are constant, and the fraction of Northern products that have not yet been imitated is constant. Furthermore, prices of Northern and Southern products and therefore, profits of Northern firms are constant. Let us choose the marginal costs of production of Northern firms as the numeraire, and set $w^N(t)b^N(t) = 1$ for all t.

First, we turn to the first-order conditions of the household's maximization problem. It follows that the optimal evolution of consumption of Northern and Southern households, i.e. the Euler equation, in steady state is given by

$$g = r - \rho \tag{3.12}$$

which implies equal growth rates in North and South. Households budget constraints in steady state are given in Appendix 3.A.3.

Now, consider the equilibrium in the labor markets. The resource constraint in the South (3.8) becomes

$$\beta L = gmF^S + mb^S L. \tag{3.13}$$

A higher fraction of products that have been imitated m implies that there is more imitation activity in the South so that on average Northern products are copied sooner, ceteris paribus. This tends to depress innovation activity in the North implying a lower g. The resource constraint of the North (3.7) can be written as follows in the steady state

$$(1-\beta)\,L = gF^N + Lb^N\,(n-m) + (1-\beta)\,Lb^N\,(1-n) \tag{3.14}$$

[13] Because of equal interest rates, consumption growth is identical across regions in steady state. Hence, net foreign assets will remain zero forever. If net foreign assets are non-zero, T_S is to be interpreted as sum of transfer and interest payments. For a formal derivation see Appendix 3.A.4.

where n denotes the "consumption gap" between South and North. Note that a higher share of the South in total production m releases resources from the production sector in the North that can be reallocated to the R&D sector, ceteris paribus. This allows the North to introduce new products at a higher rate g. Furthermore, a higher consumption share of the South n induces a reallocation from the R&D sector to the production sector in the North to satisfy the additional demand for existing Northern products by the South, thereby depressing innovation in the North, ceteris paribus.

Next, a fixed inter-sectoral allocation of labor implies that prices of Northern products must be constant in steady state. We denote the price of a new product that is sold exclusively to Northern households as z_N. Since all firms face the same demand curve and have the same cost structure, z_N is identical for all new products $j \in (N_S(t), N(t)]$. From the arbitrage condition (3.6) follows that prices for all old Northern goods $j \in (N^S(t), N_S(t)]$, which are sold to all households, are also constant and determined by $z_S = \beta + (1-\beta)\, z_N$. Moreover, this implies that profits are constant over time. Prices of Southern products $w^N(t)b^N(t)$ are equal to 1 due to our choice of numeraire. This is consistent with the steady state, else demand for Southern labor would change over time.

Let us consider the *average* life cycle in steady state of some product j, which is introduced at time t. At the time of introduction product j is sold at price z_N exclusively to Northern households. At time $t + \Delta$, where $N(t) = N_S(t+\Delta) = N_S(t)\exp(g\Delta)$, the Northern firm producing good j lowers the price to $\beta + (1-\beta)\, z_N$ and exports it to the South. Therefore, the average demand lag equals $\Delta = -\log(n)/g > 0$, decreasing in the consumption share n and the innovation rate g. We consider the case where $\widetilde{T} > t + \Delta$ for all t. In other words, on average Northern products are exported to the South for some time before they are copied by a Southern firm.[14] Notice that in steady state $\widetilde{T}$ follows an exponential distribution. Thus, the average time span product j is being produced in the North is determined by $1/\mu$. Hence, our assumption above implies that the demand lag is shorter than the time span product j is being manufactured in the North, i.e. $1/\mu > \Delta$. Due to lower production costs in the South, Southern firms can set a price marginally below 1, the marginal costs of Northern firms. Hence, the Northern firm stops producing product j and the product is now exported to the North. Of course, this discussion is only relevant for the *average* product. By the random nature of imitation there will be some products that are imitated before households in the South become rich enough to afford them. Hence, those products will skip the export stage. The average life cycle of product j in terms of sales volume is depicted in Figure 3.6 below.

From the definition of the imitation rate $\mu = \dot{N}^S(t)/N^N(t)$ we can express the production share of the South in the total number of differentiated products as

$$m = \frac{\mu}{g+\mu} \tag{3.15}$$

which must be constant in the steady state. Next, the zero-profit condition (3.9) together with

[14]Note that in the other case with $\widetilde{T} \leq t + \Delta$, goods would on average skip the export stage. We consider the case in the text to be the interesting one.

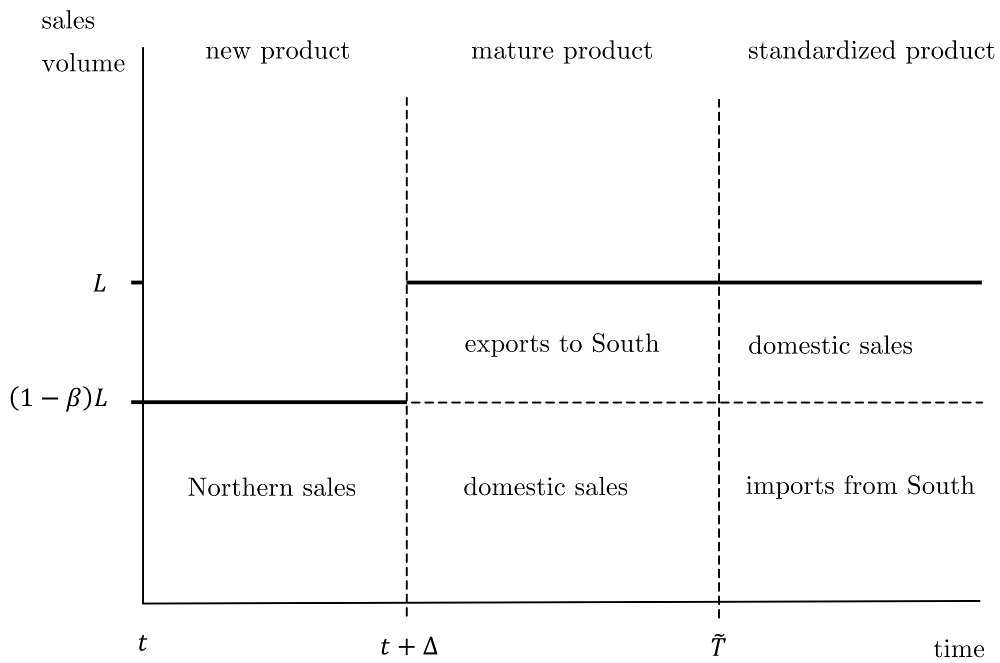


Figure 3.6: Average life cycle (in terms of sales volume)

the arbitrage condition (3.6) in the North implies that in the steady state the value of a firm is equal to the expected present discounted value of its future profits

$$\frac{[z_N - 1]\,(1-\beta)\,L}{r+\mu} = \frac{F^N}{b^N}. \tag{3.16}$$

Similarly, in the South the zero-profit condition (3.10) yields

$$\frac{\left[1-\omega^S b^S\right] L}{r} = \omega^S F^S \tag{3.17}$$

where $\omega^S(t) \equiv w^S(t)/N(t)$ is constant since wages in the South grow at rate g. Last, in steady state, the balance of payments (3.11) becomes

$$(n-m)\,[\beta + (1-\beta)z_N]\,\beta = m(1-\beta) + \beta T \tag{3.18}$$

where $T \equiv T_S(t)/N(t)$. Note that due to Walras' law the balance of payments is implied by the budget constraints, the zero-profit conditions and the resource constraints.

Equations (3.12)-(3.18) in the unknowns g, μ, n, m, r, z_N, and ω^S fully characterize the steady state. We can reduce this system to 2 equations in 2 unknowns m and g. The first equation, the *RS-curve*, describes a steady state relationship between g and m that is consistent with labor market clearing in the South:

$$m = \frac{\beta L}{g F^S + b^S L}. \tag{3.19}$$

The second equation, the *NA-curve*, describes a steady state relationship between g and m that is consistent with labor market clearing in the North, balance of payments, free entry in the North, and the no arbitrage condition

$$\left(1+\rho\frac{F^N}{b^N L}+\frac{g}{1-m}\frac{F^N}{b^N L}\right)\left((1-\beta)\left(\frac{1}{b^N}-1+m\right)-g\frac{F^N}{b^N L}\right)=m(1-\beta)+\beta T. \quad (3.20)$$

To guarantee that the *NA-curve* defined by (3.20) has a positive x-axis intercept in the (m, g) space we make the following assumption.

Assumption 3.1. $\left(1+\rho\frac{F^N}{b^N L}\right)(1-\beta)\left(\frac{1}{b^N}-1\right)\geq \beta T.$

Proposition 3.1. *Given Assumption 1 holds, a steady state equilibrium with positive growth rate g and a constant share of imitated products m exists.*

Proof. The *RS-curve* (3.19) is downward sloping in the (m, g)-space. To determine the shape of the *NA-curve* we rewrite (3.20) as $NA(m, g) = 0$. The left hand side of this equation is a quadratic function in g with inverted U-shape. If Assumption 1 holds, $NA(m, g)$ has a negative and a positive solution for g. Thus, $NA_g(m, g) < 0$ at the relevant solution. Further, differentiation shows that $NA_m(m, g) = \frac{g}{(1-m)^2}\frac{F^N}{b^N L}\left((1-\beta)\left(\frac{1}{b^N}-1+m\right)-g\frac{F^N}{b^N L}\right)+\left(\rho\frac{F^N}{b^N L}+\frac{g}{1-m}\frac{F^N}{b^N L}\right)(1-\beta)>0$. Hence, the *NA-curve* has a positive slope and positive intercept with the x-axis and a negative intercept with the y-axis in the (m, g)-space. Figure 3.7 below depicts the graphical solution of the steady state.[15] □

3.4.6 Transitional Dynamics

The transitional dynamics are easy to characterize. The full derivation of the transitional dynamics including a phase plane illustrating the dynamics is given in Appendix 3.A.2. If we replace g with g^S at the axis in Figure 3.7 above, the *RS-curve* now determined by equation (3.8), representing the Southern full employment condition only, must hold also outside the steady state. Hence, along a transition path, m and g^S move along the *RS-curve*. The *NA-curve* (3.20), instead, is a steady state condition. Appendix 3.A.2 further demonstrates that the steady state is sattle-path stable. When the number of industries in the South is below its steady state value, $m(0) < m$, then $\dot{m}/m = g^S - g > 0$, i.e. the growth rate of imitation is higher than the growth rate of innovation during the transition process. Thus, m converges monotonically to its steady state value.

[15]Note that the $m|_{g=0}$ implied by $NA|_{g=0}$ is given by $m = -\left[\left(1+\rho\frac{F^N}{b^N L}\right)(1-\beta)\left(\frac{1}{b^N}-1\right)-\beta T\right]/\rho(1-\beta)\frac{F^N}{b^N L} < 0$ due to Assumption 3.1. Furthermore, evaluating the *NA-curve* at $m \to 1$ yields

$$1+\rho\frac{F^N}{b^N L}+\frac{g}{1-m}\frac{F^N}{b^N L}=\frac{(1-\beta)+\beta T}{(1-\beta)\frac{1}{b^N}-g\frac{F^N}{b^N L}}$$

where the left-hand side goes to infinity as $m \to 1$ for any $0 < g < \infty$. Hence, the right-hand side goes to infinity as $g \to (1-\beta)L/F^N$.

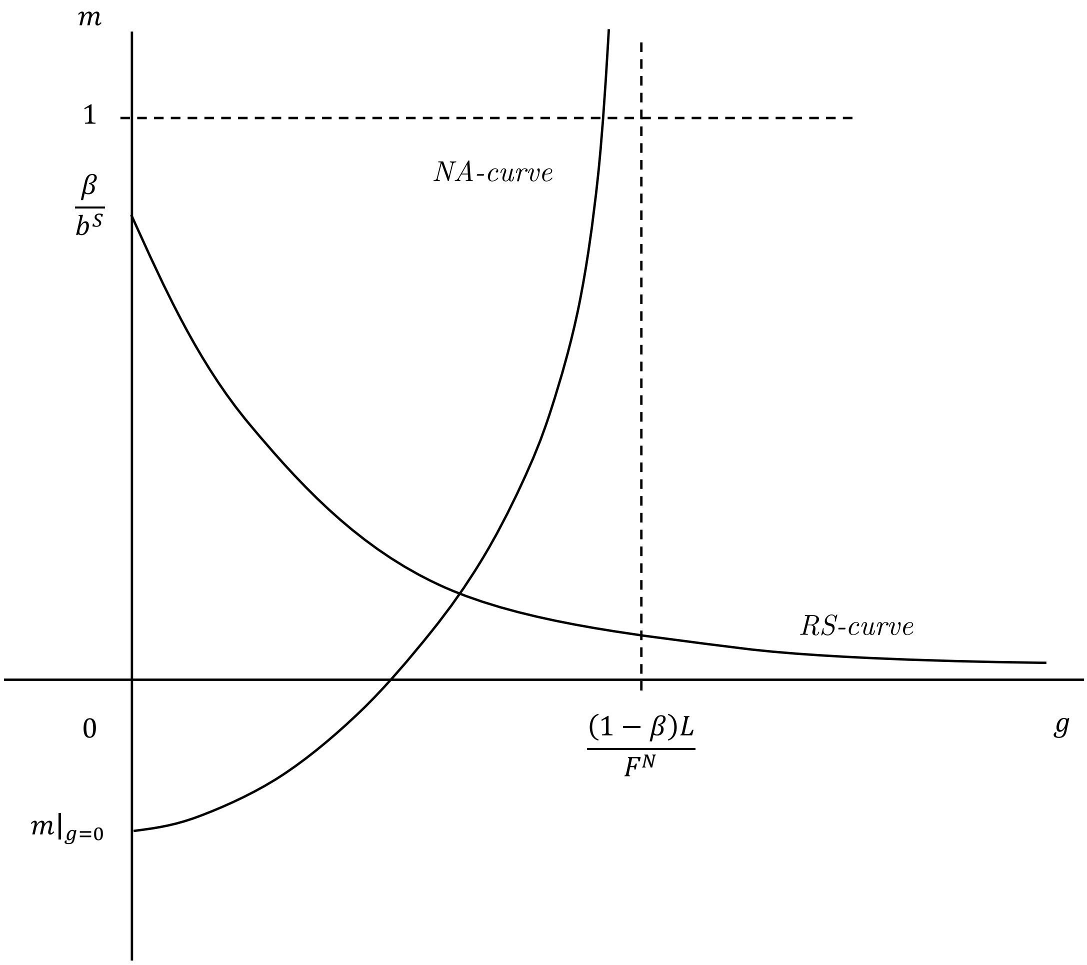


Figure 3.7: Steady state

3.5 Comparative Statics

In this section we explore the steady state implications on the growth rate and the length of the product cycle for the following changes: An increase in imitation productivity in the South and an exogenous change in inequality across regions. We simulate comparative statics results. The simulation results are depicted in Appendix 3.A.5. Note that the wage rate of North relative to South is determined by $w^N(t)/w^S(t) = \left(\omega^S b^N\right)^{-1}$, where $\omega^S = (b^S + (\rho+g)F^S/L)^{-1}$ is pinned down by (3.17). Using the zero-profit condition (3.16) and the arbitrage condition in the North, we get an expression for the terms of trade of the North, $z_S = 1 + [\rho + g/(1-m)]F^N/b^N L$.

3.5.1 Increase in Southern Labor Productivity

Proposition 3.2. *An increase in Southern productivity, i.e. a decrease in b^S or F^S, results in a higher growth rate g, Southern imitation share m and imitation rate μ. Hence, the average time span a product is being manufactured in the North $1/\mu$ becomes shorter. While the terms of trade move in favor of the North (z_S increases), two opposing effects move relative wages* *and the consumption share. Higher Southern productivity tends to increase Southern relative wages while the higher growth rate g tends to decrease them. A higher imitation rate expands the Southern consumption share whereas the higher growth rate dampens it.*

Proof. A decrease in b^S shifts the *RS-curve* upwards, whereas a decrease in F^S rotates the *RS-curve* upwards, both leaving the *NA-curve* unaffected. Hence, both a decrease in b^S and F^S lead to a higher growth rate g and Southern imitation and consumption share m. The imitation rate increases, as μ depends positively on g and m. According to the Northern zero profit condition (3.16) z_N and z_S increase. Using the Southern zero profit condition (3.17) we see that ω^S increases with higher productivity in the South but decreases in g. This implies that relative wages $w^N(t)/w^S(t)$ decrease due to the direct productivity effect and increase because of a higher growth rate. Using the Northern resource constraint, we see that a higher g reduces n while the higher m raises n. □

Intuitively, a reduction in F^S or b^S triggers more imitation because it is cheaper to produce imitated goods. An increase in μ leads to a higher risk-adjusted interest rate and hence to a lower present discounted value of profits earned from innovation, reducing the incentives to innovate. At the same time, households in the South become relatively richer (ω^S increases). Their willingness to pay increases and hence Northern terms of trade z_S improve. Moreover, they can afford to purchase a broader range of products (n rises). Higher Northern prices raises profits from developing new products and therefore, the incentives to innovate increase. This latter effect dominates such that the innovation rate g rises. However, the imitation rate μ increases more than the innovation rate g, leading to an increase in the fraction of imitated goods m in equilibrium. The increase in Southern labor productivity tends to directly increase the wage rate ω^S, holding g constant. However, there is also an indirect effect through the increase in g, which leads to an increase in the interest rate r and therefore, to a decrease in the present discounted value of profits. The indirect effect induces less firms to enter the market in the South, which tends to depress labor demand in the South and hence the wage rate ω^S. In our simulations the direct effect dominates such that the Northern relative wage rate $w^N(t)/w^S(t)$ decreases. As Southern households grow richer, not only their willingness to pay for Northern products increases but also the range of consumed goods. This increases the consumption share of the South. At the same time, a higher innovation rate dampens the consumption share. In our simulations the first effect dominates, i.e. the consumption share increases.

Effect on the Product Life-cycle

The time length Δ where products are exclusively sold in the North becomes shorter due to two reasons: households in the South are relatively richer (n rises) and the overall growth rate g is higher. Since the imitation rate μ increases, the average time span a product is being manufactured in the North $1/\mu$ becomes shorter as well. The third stage during which the North imports a product clearly increases. The time period during which the North exports a product $(1/\mu - \Delta)$ decreases according to our simulations.

Limit Case: Costless Reverse Engineering

Suppose, less realistically, that reverse engineering is costless, i.e. $F^S \to 0$. As $F^S \to 0$ the *RA-curve* rotates upwards and becomes a horizontal line at $= \beta/b^S < 1$. Since the *NA-curve* is independent of F^S, the growth rate g and the Southern imitation share m increase as $F^S \to 0$. The intuition is that if imitation becomes costless, the South can take over the maximal share of production from the North. This releases resources in the North that can be allocated to the research and development of new products. At the same time, the imitation activity is at its maximum in the South (only restricted through limited resources), which means that the risk-adjusted interest rate peaks, depressing the present discounted value of profits earned from creating new products. In equilibrium, the first effect dominates the second one, and g reaches a maximum.

3.5.2 Changes in Income Inequality across Regions

Proposition 3.3. *An increase in T, i.e. lowering world income inequality, leads to a new steady state where the growth rate g is lower and the share of imitated and consumed products, m and n, are higher. Northern relative wages deteriorate. There are opposing effects on the terms of trade and on the three stages of the product life cycle.*

Proof. An increase in T leads to an upward shift of the *NA-curve* since $NA|_{g=0}$ is a positive function of T. Notice that g implied by the *NA-curve* as $m \to 1$ is given by $g = (1-\beta)L/F^N$, independent of T. As the *RS-curve* is unaffected by a change in T, the new steady state has a lower g and higher m. Using the Northern resource constraint, we see that a lower g together with a higher m increases n. Since ω^S is a decreasing function of g, Northern relative wages $w^N(t)/w^S(t) = \left(\omega^S b^N\right)^{-1}$ are lower. A lower g tends to decrease the terms of trade z_S, whereas a higher m tends to increase them. □

A higher transfer leads to higher incomes in the South and lower incomes in the North, ceteris paribus. Lower incomes in the North depress the incentives to develop a new product, which decreases the innovation rate g. As Southern resources are fixed, the fraction of imitated products increase. At the same time, higher incomes in the South translate into a higher willingness to pay for older products produced in the North. This implies that profits of innovating firms in the North from selling only to Northern households fall short of profits from selling to all households, creating a disequilibrium in the North. This induces some Northern firms to start exporting. As Southern households consume more products, i.e. $N_S(t)$ increases, their marginal willingness to pay, ceteris paribus, decreases until the equilibrium in the North is restored. In the new equilibrium, households in the South consume a higher fraction of all products n, and their (marginal) willingness to pay is lower. In our simulations, the North's export prices z_S decrease, and as the North's import prices are equal to one, the terms of trade move in favor of the South.[16]

[16]Totally differentiating the Northern zero-profit condition (3.16) and the definition of the imitation rate (3.15) shows that $dz_N > 0$ and $d\mu < 0$ if and only if $\beta/\left(\beta - b^S m\right) > m/(1-m) > b^S L/F^S$, where $\beta/b^S > m$

Effect on the Product Life-cycle

There are two opposing effects on the the first stage of the product cycle (the demand lag Δ). On the one hand, households in the South are richer so that the Northern firm producing the latest product would like to export sooner (effect of higher n). On the other hand, even though the level of income for Southern households is higher, their income grows at a lower rate. This induces the Northern producer of the latest product to export later (effect of lower g). The simulations show that the first effect dominates so that the first stage, where new products are exclusively sold in the North, becomes shorter. There are two opposing effects on the second stage of the product cycle. On the one hand, the imitation rate μ decreases because of a lower growth rate g. On the other hand, the higher share of imitation increases μ. In our simulations the effect of a lower growth rate dominates. Hence, the average time span a product is being manufactured in the North $1/\mu$ becomes longer so that the third stage during which the North imports a product decreases. Moreover, the time period during which the North exports a product $(1/\mu - \Delta)$ becomes longer.

3.5.3 Comparison to CES Utility Case

An important advantage of the non-homothetic utility function assumption is the possibility to analyze the effect of the demand side on the average time span of the three product cycle stages. With CES preferences there is no first stage where the product is exclusively produced and consumed in the North as all consumers consume all goods. This is a main difference of this chapter compared to Grossman and Helpman (1991a). Besides being able to discuss effects on the first stage, the model presented in this paper is similar to the narrow-gap case in Grossman and Helpman (1991a). In the following, we are going to elaborate some differences and similarities in the comparative statics results.

A change in Southern productivity has similar effects. An improvement in Southern productivity in imitation leads in both models to higher growth and imitation rates. Relative wage rates move in favor of the South. In Grossman and Helpman (1991a) the Northern terms of trade are connected to the change in relative wage rates (prices are a constant markup over marginal costs), and hence deteriorate. In this paper, the terms of trade depend on the willingness to pay of households and hence move in favor of the North. As Southern relative wages increase, Southern households' willingness to pay for Northern products increase.

Next, we have a closer look at changes in inequality. In Grossman and Helpman (1991a), a transfer from rich North to poor South has no effects on innovation incentives because of CES preferences. Instead, let us compare country size effects for which we provide simulation results in Appendix 3.A.5. In Grossman and Helpman (1991a), an increase of the Southern labor force leads to an acceleration of innovation and imitation. Moreover, relative Southern wage rates increase. In our paper, the effect of a larger Southern population share leads also to a higher imitation rate. The innovation rate, however, decreases if South is not too productive in producing and imitating. This can be easily seen using the two equilibrium conditions, the

---

and we used that along the *RS-curve* $dm/dg < 0$. Sufficient conditions are $m < 0.5$ and $b^S L/F^S < 1$.

RS-curve and *NA-curve*. Both curves shift upwards with an increase in β. However, the *RS-curve* shifts up more than the *NA-curve*, and hence the positive effect on g becomes larger, the smaller b^S is. A higher population share in the South implies a higher production share in the South which releases resources in the North for innovation. This channel tends to dominate with a higher productivity in the production sector of the South. Northern relative wage rates decrease here, too. Northern terms of trade, however, increase as the Southern willingness to pay for Northern products increases with the rise in relative wages.

3.6 Extensions

Due to our assumption of symmetric preferences and identical cost structures the product cycle of product j is indeterminate. In order to show that the product cycle we impose in our baseline model emerges from more complex models, without changing the basic channels through which the income distribution operates, we discuss two extensions. It is straightforward to either change the assumptions about preferences or about technology such that the indeterminacy vanishes.

3.6.1 Hierarchic Preferences

Following Foellmi and Zweimüller (2006), we assume that households have the following non-homothetic preferences

$$u(c(j,t)) = \int_{j=0}^{N(t)} j^{-\gamma} c(j,t) dj$$

where the parameter $\gamma \in (0,1)$ determines the "steepness" of the hierarchy, i.e. how fast marginal utility falls in the index j. One can view low-indexed products as satisfying more basic needs relative to higher-indexed products. It is straightforward to derive the willingness to pay for good j, which is given by $z(j,t) \equiv j^{-\gamma}[u(\cdot)\lambda(t)]^{-1}$, and decreases in the index j. In other words, households demand, and therefore Northern firms develop, products along the hierarchy, starting with low-indexed products and gradually moving up the hierarchy ladder. This implies that profit-maximizing prices for Northern products, and hence profits decrease in the index j, given all firms have the same cost structure.

We continue to assume that Southern households can afford to consume some products manufactured in the North. Which Northern firms do not export and which firms do? First, suppose that no firm in the North exports. In that case Southern households would not exhaust their budget constraints and their willingness to pay would become infinitely large. This implies that prices for the lowest-indexed products, which have not yet been imitated by the South, become infinitely high. Hence, the firms producing the lowest-indexed products have an incentive to start exporting their products. Second, consider the case where all Northern firms export. In that case, Northern households would not exhaust their budget constraints, and their willingness to pay for an additional product would become infinitely high. This implies that new firms enter the market along the consumption hierarchy, manufacturing products that Southern households cannot afford, and that are therefore not exported.

We keep our assumptions about technology in the North. However, instead of assuming that Southern firms target Northern products at random for imitation, we assume that patents expire at time $\widetilde{T} < \infty$, where $\widetilde{T}$ is now deterministic. Random imitation would imply that there might be "holes" in the hierarchy of products. Southern firms must still invest a fixed amount of resources, e.g. building local production facilities, reverse engineering or learning the production process, in order to manufacture products, whose patents have expired, at constant marginal costs. The fixed cost implies that it is never a dominant strategy for a Southern firm to copy a product, which has already been imitated by another Southern firm. After the patent expires the Southern firm imitating the product enters into price competition with the original Northern innovator, which leads to a limit price equal to marginal costs of the Northern firm, and the exit of the Northern firm.

In sum, this model would generate the following deterministic product cycle in a steady state. At some time $t \geq 0$, the Northern firm j introduces the lowest-indexed product that has not yet been invented. It starts selling its product to Northern households at the price $z_N(j,t)$ since only they can afford to purchase new products that satisfy relatively non-essential needs. The price $z_N(j,t)$ increases at rate γg until after Δ periods, which is still determined by $N(t) = N_S(t)\exp(g\Delta)$, the Northern firm finds it attractive to lower the price to $z_S(j,t)$ and starts exporting its product.[17] The price $z_S(j,t)$ increases at rate γg until after $\widetilde{T} > t + \Delta$ periods the patent expires, a Southern firm copies the product and price competition drives the Northern firm out of the market. The price drops to the marginal cost of production of Northern firms, and stays constant from then on. Hence, such a model would eliminate the indeterminacy of the product cycle. However, the analysis would be substantially more complicated without presumably adding much additional insight.

3.6.2 Learning-by-doing

In the following, we keep our assumptions from the basic model about preferences (Section 3.4.2) and technology in South (Section 3.4.3). However, we follow Matsuyama (2002) and assume that there is passive learning-by-doing (i.e. externality of the manufacturing process) in the production sector of the North. In particular, we assume that producing one unit of output requires $b^N(j,t) = b^N\left(Q(j,t)\right)/N(t)$ units of labor, where $b^N(\cdot)$ is a decreasing function of the discounted cumulative output determined by

$$Q(j,t) = \delta \int_{-\infty}^{t} C(j,s)\exp\left(\delta\left(s-t\right)\right)ds$$

[17] This follows from taking the time derivate of the willingness to pay for the most recently innovated product $N(t)$, which is given in the steady state by $\dot{z}_N(N(t),t)/z_N(N(t),t) = r - \rho - g$. In a steady state where the allocation of resources in the North is constant across sectors the price of the newest product must be constant, i.e. $r = \rho + g$. In the steady state $n = N_S(t)/N(t)$ must be constant too, so that the price of any product j evolves over time as follows $\dot{z}_i(j,t)/z_i(j,t) = r - \rho - (1-\gamma)g$ for $i \in \{N, S\}$. Hence, using $r = \rho + g$ yields $\dot{z}_i(j,t)/z_i(j,t) = \gamma g$. Note that the firm selling the newest product must be indifferent in equilibrium whether to export or not, i.e. $\left[z_N\left(N_S(t),t\right)-1\right](1-\beta) = \left[z_S\left(N_S(t),t\right)-1\right]$, where $z_N\left(N_S(t),t\right) = n^{-\gamma} z_N\left(N(t),t\right)$.

where $\delta > 0$ can be interpreted as both the speed of learning as well as the rate of depreciation of the learning experience. Again, $C(j,t) \in \{0, (1-\beta)L, L\}$ denotes market demand. Due to depreciation the cumulative learning experience $Q(j,t)$ is bounded from above by $C(j,t)$, and can therefore not exceed L. We continue to assume that the creation of a new product requires $F^N(t)$ units of labor. As in the previous section, we assume that patents expire after $\widetilde{T} < \infty$ periods.

Again, consider a situation where Southern households can afford to purchase some of the products made in the North. Prices of Northern and Southern products are still determined as before. Our assumptions about technology imply that profits of Northern firms increase with production experience, ceteris paribus. In other words, firms which have been in the market for a longer time earn higher profits since their marginal costs are lower. In equilibrium, at any point in time some firms export and some sell exclusively to Northern households. Hence, there must be some threshold value $Q(N_S(t), t)$, implicitly defined by

$$\left[z_N(N_S(t), t) - w^N(t)\frac{b^N(Q(N_S(t), t))}{N(t)}\right](1-\beta) = \left[z_S(N_S(t), t) - w^N(t)\frac{b^N(Q(N_S(t), t))}{N(t)}\right],$$

at which a Northern firm is indifferent between exporting or not. Below this threshold value the profits from excluding Southern households exceed the profits from exporting, and vice versa. In other words, below the threshold value $Q(N_S(t), t)$ the price effect dominates the market size effect, and vice versa.

Hence, this model would imply that products go through the following cycle in steady state. A new product introduced by a Northern firm is first sold at high prices z_N only in the domestic market since this firm has a relatively low productivity level at which the price effect dominates the market size effect. The Northern firm finds it optimal to lower the price to z_S and start exporting its product after Δ periods (still determined as before) since incomes in the South grow and the Northern firm becomes more productive. At time $\widetilde{T} > t + \Delta$ the patent of the product expires, and it is imitated by a Southern firm. Price competition implies that the limit price drops to marginal costs of Northern firms, and the Northern firm exits the market. From then onwards the product is imported by the North from the South.

3.7 Conclusion

Vernon's (1966) celebrated product cycle theory hypothesizes that new products go through the following stages. In the first stage, new products are developed and introduced in high-income countries. Later in the cycle, incomes in the poorer countries have grown sufficiently such that demand for these products appears there. Thus, products that were only consumed in high-income countries before are now exported. In the third stage, production moves from high-income countries to low-income countries because they have learned the technology to produce these goods and are able to produce them at lower costs.

In this chapter we provide suggestive evidence for the different stages of the product cycle hypothesis. We show that 6 major consumer durables appear to have gone (or still go) through a "typical" product cycle. In particular, new products are not introduced simultaneously across countries and the lag in introduction depends negatively on relative GDP per capita, i.e. relative to the first country where a product is introduced. In other words, new products are introduced in affluent countries before they are introduced in less prosperous countries.

The chapter contributes to the literature by building a dynamic general-equilibrium model that is able to generate the three stages of the product cycle described by Vernon (1966). In this model, a wealthy North develops new products, which a poor South randomly attempts to copy. Besides technology, the incentives to innovate and imitate are determined by the distribution of income across regions. In other words, the demand side is an important determinant of the product cycle stages. Aside from analyzing changes in Southern labor productivity, we elaborate the effects of a redistribution of income between North and South such that inequality across regions decreases. We show that a decrease in inequality across regions leads to a decline in the innovation rate and hence a slowdown of imitation activity in the South (for a given share of the South in total production). Since Southern households are wealthier after the redistribution of income, they can afford to purchase a higher share of goods available in the world market (in particular, they can afford more newer goods produced in the North). Since Southern households are wealthier (even though their incomes grow at a lower rate), firms in the North want to export their products sooner. Therefore, the first stage of the cycle becomes shorter. At the same time the average duration new products are manufactured in the North increases because imitation activity in the South has slowed down. Firms in the South master the technology to produce a good later so that on average it takes longer for the production to move there (because of the cost advantage). Hence, the second stage of the product cycle where new goods are exported by the North to the South becomes longer. Therefore, the third stage of the cycle where the products are imitated and exported to the North becomes shorter.

3.A Appendix

3.A.1 Household Problem

Households maximize logarithmic intertemporal utility, where consumption $c\,(j,t)$ is its control, and asset holdings $a(t)$ its (endogenous) state variable

$$\max_{\{c(j,t)\}_{t=0}^{\infty}} U(0) = \int_0^{\infty} \exp\left(-\rho t\right) \log u\left(\{c\,(j,t)\}_{j=0}^{N(t)}\right) dt$$

where $\rho > 0$ denotes the time preference rate, subject to the non-negativity constraint $c(j,t) \geq 0$, and the flow budget constraint

$$\dot{a}(t) = r(t)a(t) + w(t) + T(t) - e(t)$$

with $a(0) \geq 0$, $c\,(j,t) \in \{0,1\}$, and $e(t) = \int_0^{N(t)} p(j,t)c(j,t)dj$. Furthermore, households face a no-Ponzi game condition of the following form

$$\lim_{t\to\infty} \exp\left(-\int_0^t r(s)ds\right) a(t) = 0$$

where $r(t)$ denotes the risk-free interest rate.

Households take the time paths of the interest rate, the wage rate, prices for all goods j, as well as the set of differentiated products in the economy $\{r(t), w(t), p(j,t), N(t)\}_{t=0}^{\infty}$ as given.

The current value Hamiltonian is given by

$$H\left(t, c(j), a, \lambda, \mu\right) = \log u\left(\cdot\right) + \lambda(t)\left[r(t)a(t) + w(t) + T(t) - e(t)\right] + \sum_{j}^{\infty} \xi(j,t)c(j,t)$$

where $\lambda(t)$ denotes the costate variable on the flow budget constraint and $\xi(j,t)$ the one on the non-negativity constraints. The maximum principle conditions are

$$\max_{\{c(j,t)\}_{j=0}^{N(t)}} H\left(t, c(j), a, \lambda, \mu\right) \quad \text{for all } t \in [0,\infty]\,, j \in [0, N(t)]:$$

$$\begin{aligned} u\left(\cdot\right)^{-1} - \lambda(t)p(j,t) &= 0, \quad c\,(j,t) = 1 \\ u\left(\cdot\right)^{-1} - \lambda(t)p(j,t) &\leq 0, \quad c\,(j,t) = 0 \end{aligned}$$

$$\begin{aligned} \lambda(t)r(t) &= -\dot{\lambda}(t) + \rho\lambda(t) \\ \dot{a}(t) &= r(t)a(t) + w(t) + T(t) - e(t) \\ \lim_{t\to\infty} \exp\left(-\rho t\right)\lambda(t)a(t) &= 0. \end{aligned}$$

3.A.2 Derivation of Transitional Dynamics

Using the resource constraint of the South, the relationship between g and g^S, the resource constraint of the North to substitute for g, and the balance of payments to substitute for n (assuming that it is balanced period by period), we obtain the $\dot{m}$ - schedule

$$\frac{\dot{m}}{m} = \left(\frac{1}{F^N/b^N L}\right)\left\{\frac{\lambda\beta z_N}{\beta+(1-\beta)z_N} - \left[(1-\beta)\left(\frac{1}{b^N}-1\right)+m-F\left(\frac{\beta/b^S}{m}-1\right)\right]\right\}$$

where $\lambda = \lambda_N/\lambda_S$ which is constant and equal to its steady state value, and $F = \frac{F^N/b^N L}{F^S/b^S L}$. The $\dot{m} = 0$ locus is determined by

$$\frac{\beta\lambda z_N}{\beta+(1-\beta)z_N} = (1-\beta)\left(\frac{1}{b^N}-1\right)+m-F\left(\frac{\beta/b^S}{m}-1\right).$$

It is straightforward to show that $dz_N/dm > 0$, $z_N(m) \to -\infty$ as $m \to 0$, and $z_N(m)$ equals a positive constant larger than one as $m \to \beta/b^S$ if and only if $(1-\beta)\left(1/b^N-1\right)+\beta/b^S > \lambda$ (this simply requires inequality between North and South to be sufficiently high). Thus, the $\dot{m} = 0$ locus is increasing in the (z_N, m)-space. We deduce the following dynamics for m:

$$\frac{\dot{m}}{m}\begin{cases} > 0, & z_N < z_N^* \\ = 0, & z_N = z_N^* \\ < 0, & z_N > z_N^* \end{cases}$$

where z_N^* denotes the steady state value.

The $\dot{z}_N$ - schedule is obtained by using the balance of payments, the Northern and Southern resource constraints, the definition of the hazard rate, the Euler equation, and the North's zero-profit condition

$$\begin{aligned}\frac{\dot{z}_N}{z_N} &= \left(\frac{1}{F^N/b^N L}\right)\left\{(z_N-1)(1-\beta)+\frac{\lambda\beta z_N}{\beta+(1-\beta)z_N}\right. \\ &- \left.\left[(1-\beta)\left(\frac{1}{b^N}-1\right)+m+\frac{\rho F^N}{b^N L}+F\left(\frac{m}{1-m}\right)\left(\frac{\beta/b^S}{m}-1\right)\right]\right\}.\end{aligned}$$

The $\dot{z}_N = 0$ locus is determined by

$$\begin{aligned}(1-\beta)(z_N-1)+\frac{\beta\lambda z_N}{\beta+(1-\beta)z_N} &= (1-\beta)\left(\frac{1}{b^N}-1\right) \\ &+ m+\frac{\rho F^N}{b^N L}+F\left(\frac{m}{1-m}\right)\left(\frac{\beta/b^S}{m}-1\right).\end{aligned}$$

The slope of the $\dot{z}_N = 0$ locus is given by

$$\frac{dz_N}{dm} = \frac{[\beta+(1-\beta)z_N]^2}{\beta^2\lambda+(1-\beta)[\beta+(1-\beta)z_N]^2}\left[1-\frac{F}{(1-m)^2}(1-\beta/b^S)\right].$$

We define $\widetilde{m} \equiv 1 - \sqrt{F(1-\beta/b^S)} > 0$ with $\beta/b^S < 1$, and where $\widetilde{m} > 0$ requires that $(1-\beta/b^S)^{-1} > F$, which holds e.g. in the case of identical technology, i.e. $F = 1$. It follows that $dz_N/dm > 0$ if $m < \widetilde{m}$, and vice versa. In other words, the $\dot{z}_N = 0$ locus is decreasing for $m \in \left(\widetilde{m}, \beta/b^S\right)$, and increasing for $m \in (0, \widetilde{m})$ in the (z_N, m)-space. We note that $z_N(m) \to -\infty$ as $m \to 1$ and $z_N(m)$ converges to a constant as $m \to 0$. Eventually, it follows that the dynamics for z_N are given by

$$\frac{\dot{z}_N}{z_N} \begin{cases} > 0, & z_N > z_N^* \\ = 0, & z_N = z_N^* \\ < 0, & z_N < z_N^*. \end{cases}$$

Hence, we have a system of two differential equations in m (state variable) and z_N (choice variable), whose solution is saddle-path stable. Figure A.1 below shows the phase diagram. We see that if m is below (above) its steady state value m^* it converges monotonically towards the steady state along the saddle-path.

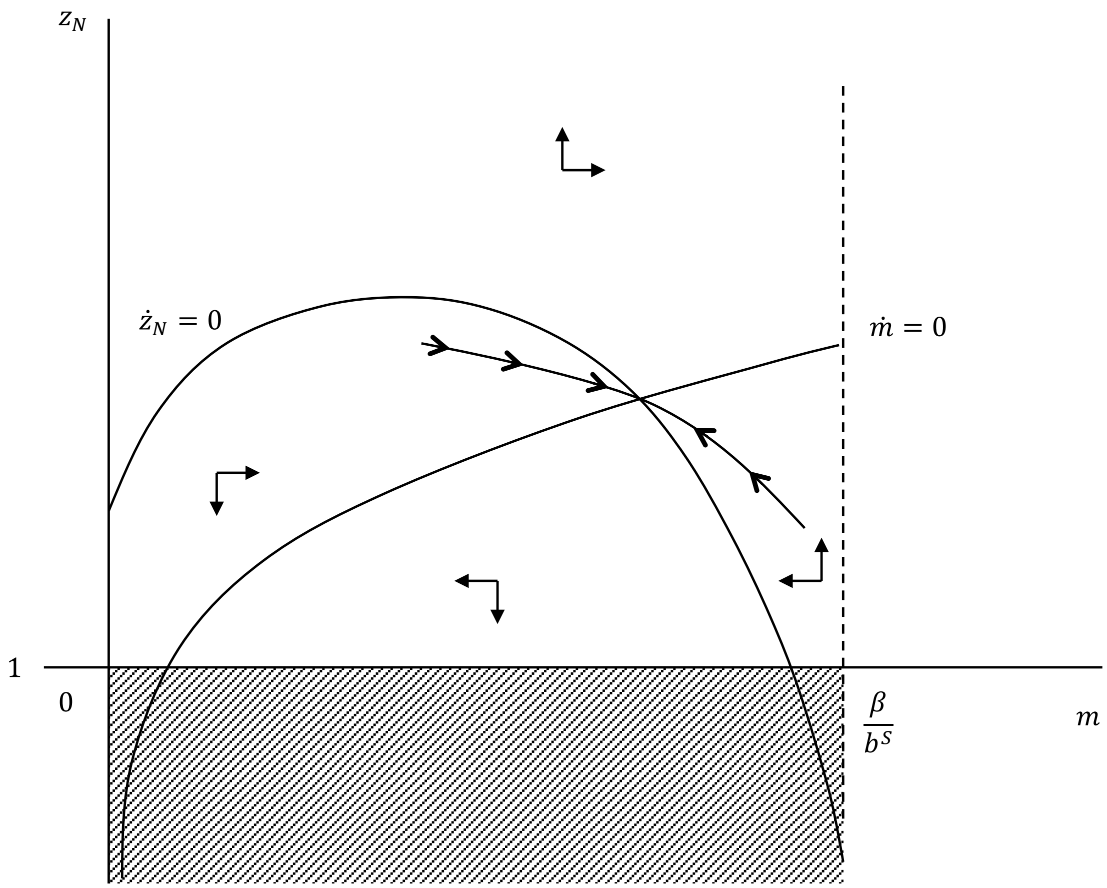


Figure A.1: Phase diagram

3.A.3 Budget Constraints

The intertemporal budget constraint of households in the North is in the steady state given by

$$N(t)\left\{m+(n-m)\left[\beta+(1-\beta)\,z_N\right]+(1-n)\,z_N\right\}=(r-g)a_N(t)+w^N(t)-T_N(t)$$

where $y_N(t)=a_N(t)+w^N(t)/(r-g)-T_N(t)/(r-g)$ denotes the lifetime income of a Northern household. We observe that Northern households save only out of their capital income (note that $r-g=\rho$), and consume all their labor income (and possible transfer income). In other words, the marginal propensity to consume out of labor and transfer income is one. Similarly, in the steady state the intertemporal budget constraint of households in the South becomes

$$N(t)\left\{m+(n-m)\left[\beta+(1-\beta)\,z_N\right]\right\}=(r-g)a_S(t)+w^S(t)+T_S(t)$$

where $y_S(t)=a_S(t)+w^S(t)/(r-g)+T_S(t)/(r-g)$ denotes the lifetime income of a household in the South. Similarly to Northern households, Southern households save only out of capital income and consume all labor income. Hence, relative lifetime incomes per capita in the steady state are (endogenously) determined by

$$\frac{y_S(t)}{y_N(t)}=\frac{\rho a_S(t)+w^S(t)+T_S(t)}{\rho a_N(t)+w^N(t)-T_N(t)}.$$

Note that in the simulations we measure inequality with the Gini coefficient. If the South's income share equals its population share the Lorenz curve lies on the 45 degree line of perfect equality, and the Gini coefficient is zero.

3.A.4 Balance of Payments

The intertemporal budget constraint of households in the South, the resource constraint in the South, and the zero-profit condition in the South imply the balance of payments as stated in the text. Due to Walras' law, the intertemporal budget constraint of the North is redundant. We drop the time index t where no confusion arises. The balance of payments in present value form at $t=0$ is given by

$$\begin{aligned}
0 &= \left\{\int_0^\infty\left[(1-\beta)LN^S\omega^N b^N-\beta L\left(N_S-N^S\right)z_S\right]\exp\left(-\int_0^t r(s)ds\right)dt\right\}\\
&+ \int_0^\infty \beta LT_S\exp\left(-\int_0^t r(s)ds\right)dt\\
&+ \left\{\beta La_S(0)-\int_0^\infty N^S\left[\pi^S-g^S v^S\right]\exp\left(-\int_0^t r(s)ds\right)dt\right\}
\end{aligned}$$

where we used $\beta LN=\dot{N}^S F^S+N^S b^S L$ from the resource constraint, $v^S=\omega^S F^S$ from the zero-profit condition, and a no-Ponzi game condition. The first two lines denote the current account, which consists of the trade balance and net transfer payments. The third line denotes net foreign asset holdings. In the steady state, we have that r and π^S are constant, N^S grows at a constant rate $g^S = g$, and $v^S = \pi^S/r$. This implies that net foreign assets become $\{\beta La_S(0) - N^S(t)\pi^S/r\}$. Hence, the balance of payments in the steady state is determined by

$$0 = \{N^S(t)(1-\beta)L\omega^N b^N - \left(N_S(t) - N^S(t)\right) z_S(t)\beta L\} + \beta LT_S(t) + \{\beta La_S(t) - N^S(t)\pi^S/r\}$$

which holds for all t in steady state, in particular at $t = 0$. Hence, it becomes obvious that if we assume initial wealth at time $t = 0$ of households in the South $\beta La_S(t)$ to be exactly equal to the present discounted value of aggregate firm profits in the South $N^S(t)v^S(t)$, net foreign assets will remain zero in steady state. We see that if Southern households would inherit sufficiently large asset holdings they could run a permanent trade deficit (even in the absence of transfers from North).

3.A.5 Simulations

We choose the following parameter configuration for our baseline simulation: $L = 1$, $F^N = F^S = 5$, $b^N = b^S = 0.75$, $\beta = 0.5$, $\rho = 0.04$, and $T = 0$.

Increase in Southern Labor Productivity

Figures A.2-A.4 show the comparative statics results of a change in labor productivity in production in the South.

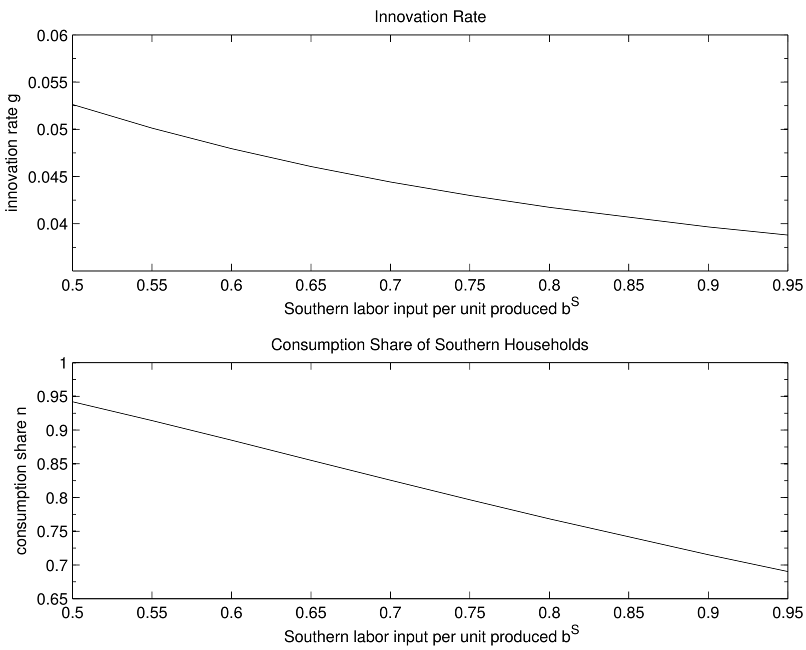


Figure A.2: Effect on innovation rate and consumption share of the South

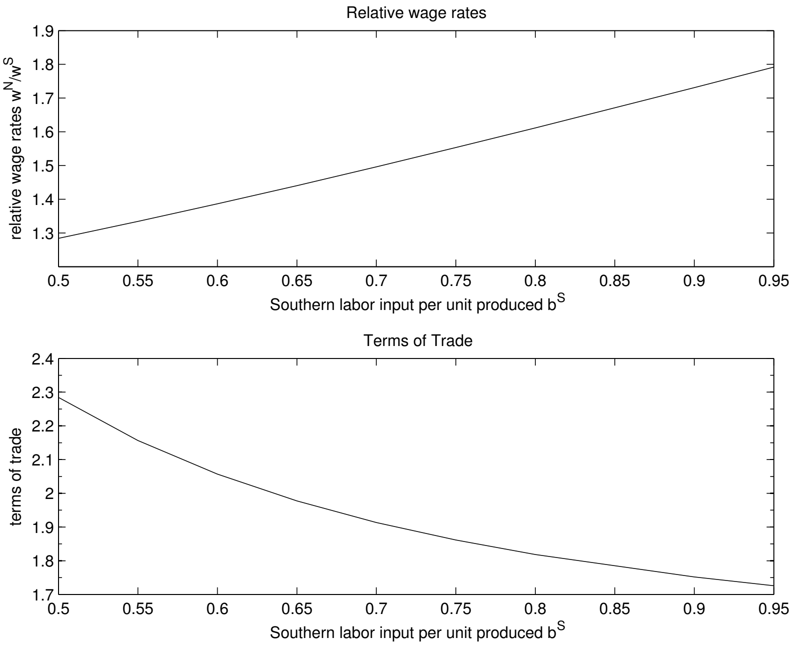


Figure A.3: Effect on relative wages

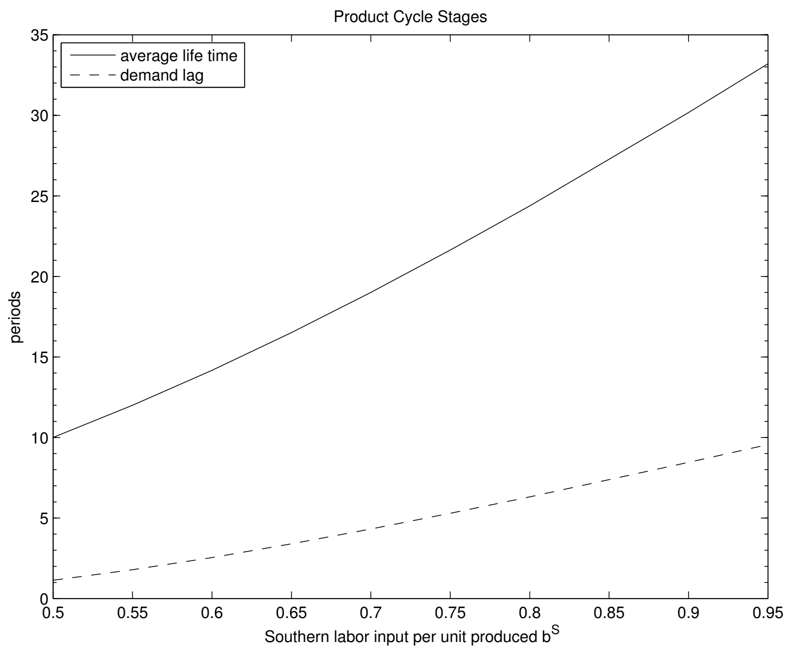


Figure A.4: Effect on the stages of the product cycle

Figures A.5-A.7 show the comparative statics results of a change in labor productivity in R&D in the South.

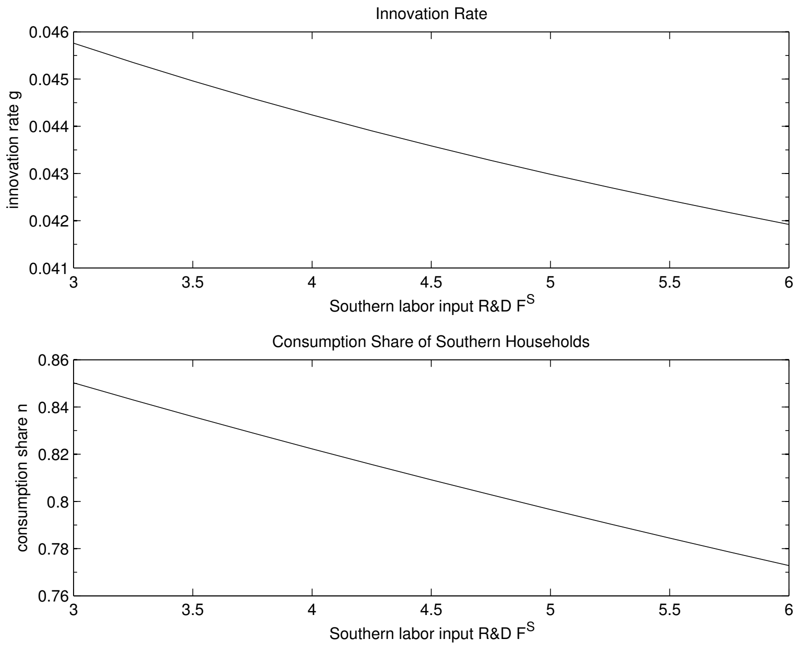


Figure A.5: Effect on innovation rate and consumption share of the South

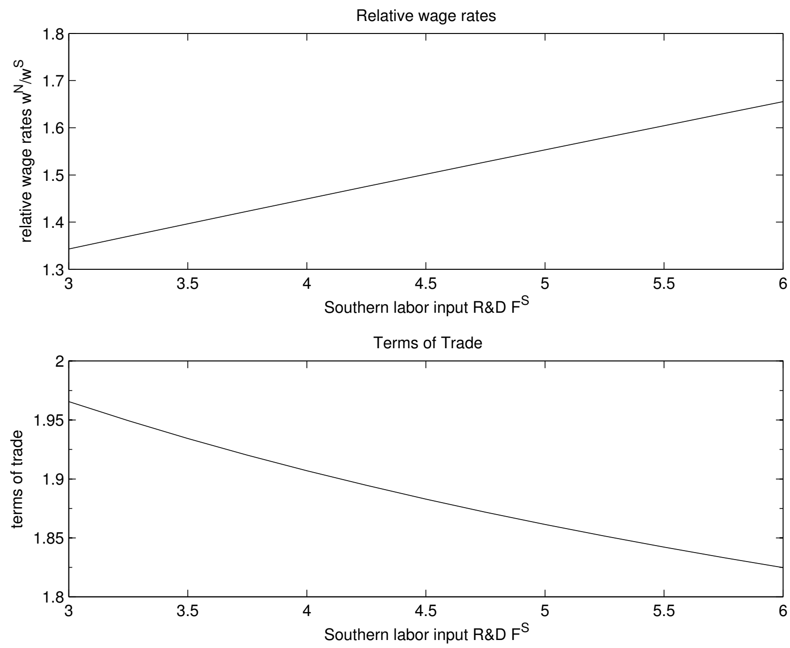


Figure A.6: Effect on relative wages

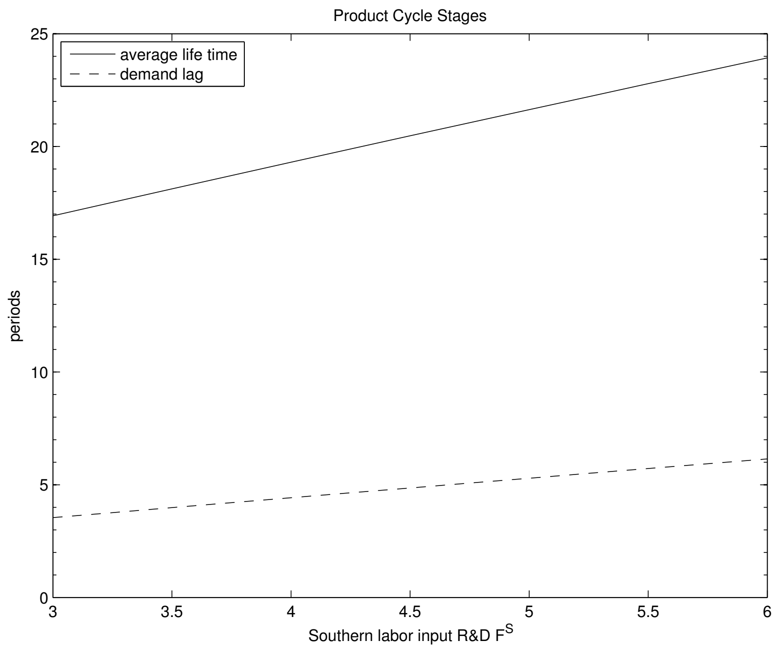


Figure A.7: Effect on the stages of the product cycle

Changes in Inequality across Regions

Figures A.8-A.10 depict the effects of an increase in inequality across regions due to a regressive transfer, i.e. a transfer from poor South to rich North.

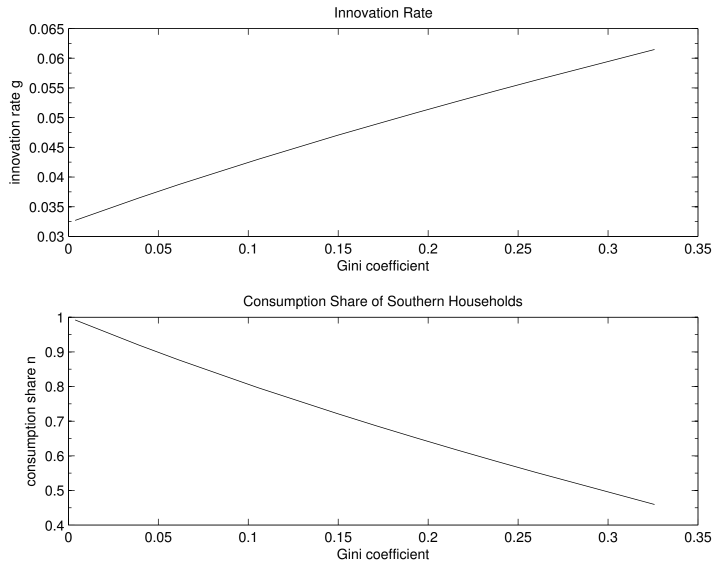


Figure A.8: Effect on innovation rate and consumption share of the South

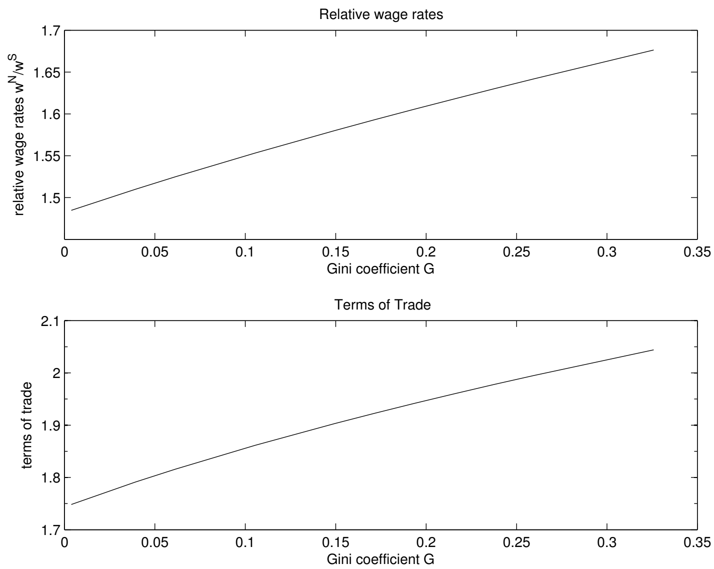


Figure A.9: Effect on relative wage rate and terms of trade

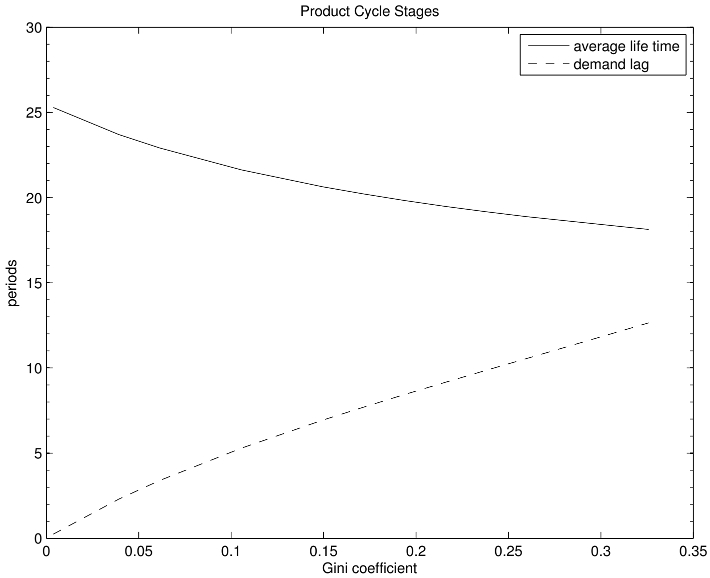


Figure A.10: Effect on stages of the product cycle

Changes in Population Shares

Figures A.11-A.13 show the effects of an increase in Southern population share.

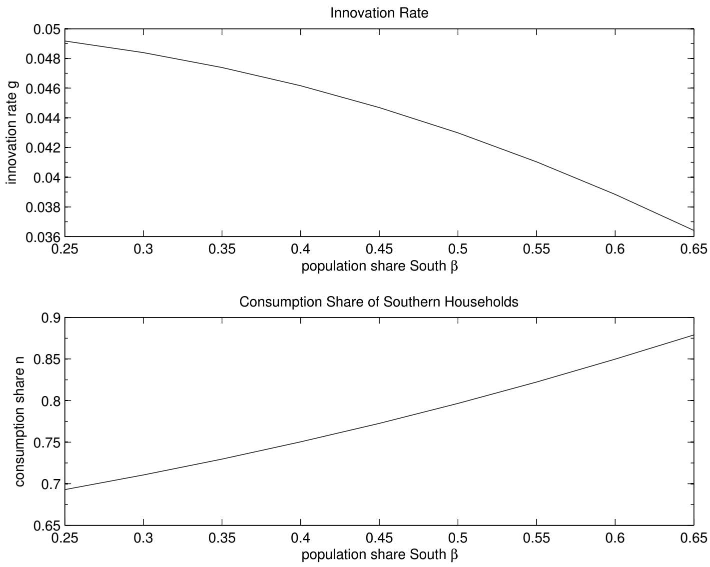


Figure A.11: Effect on innovation rate and consumption share of the South

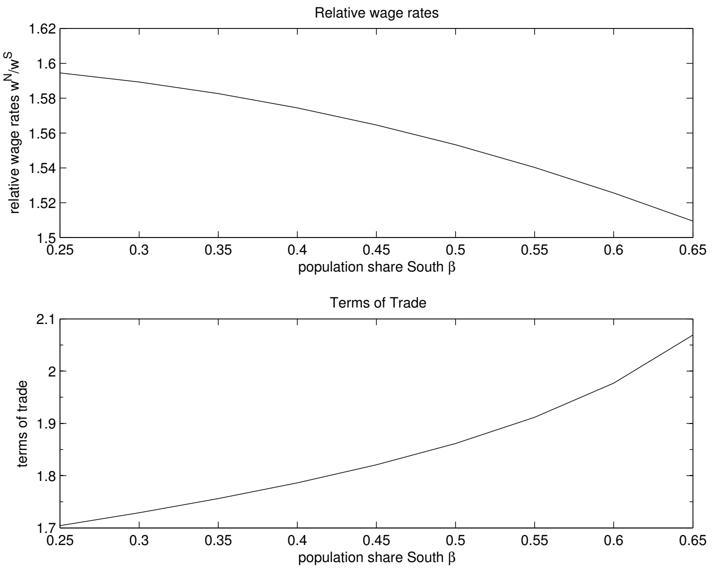


Figure A.12: Effect on relative wage rate and terms of trade

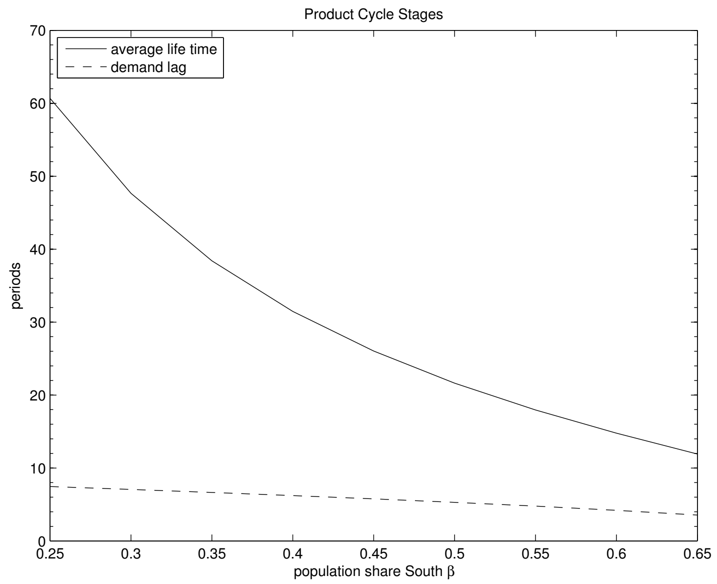


Figure A.13: Effect on stages of the product cycle

4 Income Distribution, Market Size, and Foreign Direct Investment

4.1 Introduction

> Picture Mom, Dad, and the kids in an upper-middle-class Asian family in 10 years' time: After loading up with cash at the corner Citibank, they drive off to Walmart and fill the trunk of their Ford with the likes of Fritos and Snickers. On the way home, they stop at the American-owned Cineplex to catch the latest Disney movie, paying with their Visa card. In the evening, after putting the kids to bed, Mom and Dad argue furiously about whether to invest in a Fidelity mutual fund or in a life insurance policy issued by American International Group (The New York Times, February 1, 1998).

Why do firms engage in foreign direct investment to serve a foreign market rather than export? The economic literature on foreign direct investment (FDI) and international trade regards the size of the market in the destination or host country to be a fundamental determinant of investment and trade flows. The size-of-market hypothesis as proposed by Balassa (1966), and later by Scaperlanda and Mauer (1969), argues that foreign direct investment will take place if the market is large enough to capture economies of scale. Typically in the literature, market size is reflected by the host country's aggregate income (see e.g. Markusen 2002; Davidson 1980). However, as the quote from the New York Times illustrates it might be the middle class in the host country that plays a major role in attracting FDI. In turn, this suggests that the distribution of income within the host country may be important in determining international investment and trade patterns. The business literature has long recognized that aggregate income might not be an adequate measure for the size of the market:

> The problem in using gross national product (or gross domestic product) is its failure to show for some countries that a large number of people have very low incomes. Hence, a seemingly sizable GNP might nevertheless represent a small market for many U.S. goods (Stobaugh 1969, p. 131).

A very early example of FDI also highlights the role of the market size. In 1867 the Singer Manufacturing Co., with headquarters in New York, opened a production facility for sewing machines, the first mass marketed (complex) consumer good, in the UK (see Godley 2001). According to Godley (2001) Singer's enterprise was driven by booming demand and lower production costs in the UK. Browsing today's business press confirms that the purchasing power and size of the middle class seem elemental aspects for investors in evaluating the attractiveness of markets. For instance, in a survey on India the consultancy Ernst & Young writes: "The fundamentals that make India attractive to investors remain intact. The high potential of the domestic market driven by an emerging middle class ... " (Ernst & Young 2012b, p. 2).[1]

Motivated by these observations, this chapter argues that the distribution of aggregate income within the host country is important for its attractiveness of horizontal FDI because it segments the market.[2] Consider a situation where firms face a proximity-concentration trade-off, i.e. they want to concentrate production to capture economies of scale, but at the same time locate their production in the proximity of their consumers due to trade costs. In the presence of a proximity-concentration trade-off the firm's foreign market entry mode depends on which market segment it serves. Firms serving a mass market in the foreign country engage in FDI, whereas firms catering exclusively to the needs of a few rich abroad tend to export. In other words, foreign direct investment will take place if the market is large enough to capture economies of scale. This is the essence of the size-of-market hypothesis. Poor households are likely to be irrelevant to a firm's decision because they can often barely afford the level of subsistence consumption, let alone consumer goods like cars etc. This implies that a country's middle class, in terms of per capita income and size, should be important in determining FDI and trade patterns.

First, we formalize the market size theory in a simple general-equilibrium model with two regions Home and Foreign, and low, middle and high income classes of consumers in both regions. We think of Home as a wealthy region relative to Foreign, e.g. the U.S. compared to the rest of the world. Consumers must satisfy a certain level of subsistence consumption in terms of food before they can spend income on horizontally differentiated consumer goods. We look at the case where consumers in the low income class in both regions cannot afford to purchase differentiated goods. Food is produced under conditions of perfect competition and sold only domestically. Monopolistic firms producing differentiated goods face a proximity-concentration trade-off. Due to the presence of iceberg trade costs they want to produce near their consumers while concentrate their production to take advantage of economies of scale. In

[1]There are numerous other examples. Ernst & Young (2012a) also emphasizes the importance of the middle class in its survey on the attractiveness of Russia. The McKinsey Global Institute (2010) makes a similar point in a report on African economies. In the World Investment Report 2012 prepared by the UNCTAD it is noted that a growing middle class in emerging markets has attracted FDI in the manufacturing and service sectors (UNCTAD 2012). Forbes Magazine (2012) writes that the consumer goods company Unilever built a factory in the Chinese city of Anhui mainly to produce products for China's growing middle class. In an article on Indonesia the Wall Street Journal (2013) reports that Toyota Motor Corp., Honda Motor Co. and Nissan Motor Co. invest several hundred million dollars to step up production at their plants in Indonesia as a response to increasing car purchases from the growing middle class.

[2]Traditionally, the literature distinguishes between horizontal and vertical FDI. The former refers to the duplication of a production facility abroad designed to serve customers in the foreign market, whereas the latter refers to the segmentation of the production process (i.e. outsourcing or offshoring). Our theory complements the literature on horizontal FDI. The motives for vertical FDI are usually explained by lower production costs abroad (see e.g. Blonigen 2005; Caves 2007). Several studies indicate that the bulk of FDI is horizontal rather than vertical (see e.g. Markusen and Maskus 2002; Ramondo et al. 2011). Recently, a literature concerned with platform FDI, where the foreign affiliate's output is sold in a third market rather than the host market, emerged (see e.g. Ekholm et al. 2007). Although important, in our analysis we will abstract from that phenomenon, as well as licensing (see e.g. Horstmann and Markusen 1996).

this setting, ex-ante identical firms choose different pricing strategies supplying different market segments (i.e. income classes) in equilibrium. Depending on their pricing strategy they opt for a different mode to supply the foreign market. In equilibrium, firms serving the mass market (i.e. middle and rich classes in both regions) engage in FDI whereas firms serving exclusively the rich classes in both regions export. For firms supplying the mass market the cost of exporting is higher than the cost of setting up a foreign production facility. Since they serve a large market, economies of scale are high enough to compensate higher fixed costs associated with FDI such that average costs are lower compared to exporting. Due to our assumptions about the distribution of income within regions, products sold in the mass market are priced according to the willingness to pay of the middle class in Foreign. This is the highest price a firm can set, if it wants to sell on the mass market. Thus, taking the perspective of firms in Home, we show that, ceteris paribus, redistributing income towards the middle class in Foreign increases the number of multinational corporations (MNCs) with a production facility in Foreign and headquarters (HQ) in Home. Expanding the size of the middle class in Foreign has, ceteris paribus, ambiguous effects on the number of MNCs with HQ in Home, depending on whether the poor or rich class contracts.

Second, we extend our model to differences in technologies across regions, and show that the results of redistributing income between classes are the same as in the baseline model. We further analyze the baseline model in the case of standard CES preferences, and argue that it is meaningless to distinguish between different market segments in that case because only aggregate income matters for the determination of FDI. Last, we show that the effects of changing the income distribution is not isomorphic to changing the skill distribution in the economy. Based on Markusen and Venables (2000), and Egger and Pfaffermayr (2005), we study a simple model where differentiated goods producers combine different skills in their production. In simulations we show that in contrast to the baseline model the relationship between per capita income of Foreign's middle class and the number of MNCs active in Home is ambiguous.

Third, as an illustration we empirically investigate the model's predictions regarding FDI activity, and leave the analysis of trade flows for future work.[3] We focus on the effect of the middle class' per capita income in the host country on FDI positions, using pooled data on outward FDI positions from 1997-2007 of OECD countries from the OECD International Direct Investment Statistics (OECD, 2012). Building on empirical work by Bernasconi (2013), who uses inequality and income data from UNU-WIDER (2008) and Heston et al. (2012) to construct empirical income distributions, we define global low, middle, and high income classes by imposing common income thresholds across all countries in the sample. Applying different

[3] In the trade literature, both theoretical and empirical, on the determinants of bilateral trade flows there has been some renewed interest in countries' similarity in per capita incomes or income distributions motivated by the famous Linder hypothesis (Burenstam-Linder 1961), see e.g. Markusen (1986), Hunter (1991), Dalgin et al. (2008), Fajgelbaum et al. (2009), Hallak (2010), Markusen (2010), Foellmi et al. (2011), Martinez-Zarzoso and Vollmer (2011), Bernasconi (2013). However, this development has mainly been restricted to the trade literature, and with the exception of Fajgelbaum et al. (2011) has not spilled over to the literature on foreign direct investment. For a brief discussion on the problem of estimating FDI flows/stocks and trade flows simultaneously see Section 4.6.4.

definitions of a global middle class used in the literature, we find a positive relationship between average income of the middle class in the host country and the amount of FDI it attracts from OECD countries, controlling in particular for host GDP and GDP per capita.

The rest of the chapter is organized as follows. The related literature is briefly reviewed in Section 4.2. Section 4.3 presents the baseline model, and looks at the effects of income redistribution on FDI activity and international trade. In Section 4.5.1, we extend the model to a North-South perspective allowing for differences in technology across regions. Sections 4.5.2 and 4.5.3 compare the baseline model to the standard model with constant-elasticity-of-substitution (CES) preferences, and a simple factor-proportions model incorporating different skill levels. In Section 4.6, we illustrate the predictions of the baseline model with regard to FDI activity using data on outward FDI stocks of OECD countries. Section 4.7 concludes.

4.2 Related Literature

This chapter is related to the theoretical and empirical literature on the determinants of FDI and international trade. Excellent surveys of the literature can be found in Agarwal (1980), and more recently, in Helpman (2006) or Caves (2007). A detailed treatment of multinational firms in general-equilibrium theory is given in Markusen (2002).[4]

Even though market size is deemed an important determinant of FDI and trade the literature has, by and large, focused on supply side explanations. Often, consumption patterns for consumers at different income levels are identical due to the assumptions on preferences. Thus, implying that the relevant market size for all firms is reflected by aggregate income regardless of the distribution of income.

For example, Brainard (1997) uses a simple model where consumers have homothetic preference and firms face a proximity-concentration trade-off to motivate an empirical assessment of the proximity-concentration hypothesis. Due to the assumption of homothetic preferences only aggregate GDP in the destination country matters, but not its distribution. Thus, in her regressions explaining the share of export sales in total sales (i.e. export plus foreign affiliate sales) Brainard (1997) includes aggregate GDP in the destination country to control for market size effects. She finds that foreign affiliate relative to export sales increase the higher trade costs are, and the lower are economies of scale at the plant relative to the firm level. Brainard (1997) concludes that the proximity-concentration hypothesis is quite robust in explaining the share of export sales in total sales.

Another example that focuses on the supply side is Helpman et al. (2004), which explores the emergence of multinational corporations in the context of heterogeneous firms based on Melitz (2003). They argue that only the most productive firms are internationally active, and of those only the most productive engage in FDI. Using data on U.S. exports and foreign

[4] Most theories on the organization of the firm in an international context are based on the OLI (Ownership, Location, Internalization) framework (see Dunning 1988, Dunning 2000). In this framework the existence of multinational firms is explained with competitive advantages due to ownership structure, location abroad, and internalization (net benefits from producing itself rather than licensing technology).

affiliate sales Helpman et al. (2004) confirm the prediction of their model that foreign affiliate sales relative to export sales are low in sectors where firm heterogeneity is high. Additionally, they find evidence consistent with the proximity-concentration hypothesis.

A notable exception is Fajgelbaum et al. (2011). They propose a Linder-type explanation for bilateral FDI activity, and stress that horizontal FDI is more likely to take place between countries with similar per capita incomes. In particular, they present a multi-country general-equilibrium model with vertically (quality) differentiated products. Preferences are such that consumers with higher incomes choose higher quality varieties, and firms face a proximity-concentration trade-off. They show that firms supply foreign markets that have a similar demand structure to their domestic market via FDI rather than via exports.

Blonigen (2005) surveys the empirical literature on FDI determinants. A large part of the literature investigates predictions on FDI decisions based on partial-equilibrium models. Only recently has the literature started to test general-equilibrium models. Most empirical models that employ modified versions of the gravity equation approximate market size by GDP. However, Blonigen (2005) points out that contrary to the trade literature the gravity equation lacks a theoretical foundation in the FDI literature. He concludes that the (empirical) literature is still in its infancy and most hypotheses are still up for grabs.[5]

4.3 Model

We propose a simple general-equilibrium model with two regions, Home and Foreign, where different market sizes interact with a proximity-concentration trade-off. Producers manufacturing horizontally differentiated products decide whether to serve a foreign market by exporting their product or whether to duplicate production in the foreign country, depending on the market segment they serve.

4.3.1 Environment

There are two regions, Home (H) and Foreign (F). The population size of region $i = \{H, F\}$ is denoted by L^i. In each region i there are three groups of consumers $k = \{1, 2, 3\}$, where group k in region i is of size $s_k^i L^i$, with $\sum_k s_k^i = 1$ for all i. Each consumer in region i and group k is endowed with θ_k^i (efficiency) units of labor, supplied inelastically. We make the following assumption.

Assumption 4.1. $\sum_k s_k^H \theta_k^H > \sum_k s_k^F \theta_k^F$.

In equilibrium average income in Home is larger than in Foreign if Assumption 4.1 holds.

[5]Blonigen (2005) also mentions that most determinants of cross-country FDI are statistically rather fragile. Blonigen and Piger (2011), using Bayesian statistics to gauge model selection, argue that traditional gravity variables like parent and host GDP and GDP per capita should be included in a regression explaining bilateral FDI positions. However, their analysis does not support the inclusion of many explanatory variables used in previous studies like legal institutions and business costs (e.g. time to start business etc.).

4.3.2 Consumers

Preferences follow Murphy et al. (1989). All consumers have the same non-homothetic preferences defined over a homogeneous good x (e.g. food), and a continuum of indivisible and horizontally differentiated products indexed by $j \in \mathcal{J}$

$$U = \begin{cases} x, & x \leq \overline{x} \\ \overline{x} + \int_{\mathcal{J}} c(j)dj, & x > \overline{x} \end{cases} \tag{4.1}$$

where $\overline{x}$ denotes a level of subsistence consumption of the homogeneous good that must be satisfied before consumers can start purchasing differentiated products, and $c(j)$ is equal to one if product j is purchased and zero otherwise. The homogeneous good x is a necessity in the sense that consumers' propensity to spend is one at low levels of income, and zero after the subsistence amount $\overline{x}$ is purchased. Differentiated products enter the utility of consumers symmetrically, i.e. no product is intrinsically better or worse than any other product. Indivisibility of differentiated products combined with local satiation after one unit has been purchased, implies that consumers choose their consumption only along the extensive margin but have no choice about the intensive margin. This contrasts with standard constant-elasticity-of-substitution (CES) preferences where consumers choose only along the intensive margin but have no choice about the extensive margin (see Appendix 4.A.3). Consumers maximize utility (4.1) subject to their budget constraints

$$y = p_x x + \int_{\mathcal{J}} p(j)c(j)dj \tag{4.2}$$

where $y \equiv w\theta + v$ denotes income, which consists of wage income $w\theta$ and shares in producer profits v, which will be zero in an equilibrium with free entry; p_x is the price of food, and $p(j)$ is the price of differentiated product j. The first-order conditions to the consumer's maximization problem for $x \leq \overline{x}$ are given by

$$\begin{aligned} 1 - \lambda p_x &= 0 \\ y - p_x x &= 0 \end{aligned}$$

and for $x > \overline{x}$ by

$$\begin{aligned} 1 - \lambda p(j) &\geq 0, \;\; c(j) = 1 \\ 1 - \lambda p(j) &< 0, \;\; c(j) = 0 \\ y - p_x x - \int_{\mathcal{J}} p(j)c(j)dj &= 0 \end{aligned}$$

where λ denotes the Lagrange multiplier, which can be interpreted as the marginal utility of income. Affluent consumers have a low marginal utility of income (low λ), whereas low-income consumers have a high marginal utility of income (high λ). The first-order conditions implicitly define an endogenous income threshold $\overline{y} \equiv p_x\overline{x}$. Consumers with income above $\overline{y}$ can afford to buy differentiated products whereas below they cannot. From the first-order conditions we can deduce individual demand for the homogeneous good and differentiated products. It follows that individual demand for food is given by

$$x_k^i = \begin{cases} y_k^i/p_x^i, & y_k^i \leq \overline{y}^i \\ \overline{x}, & y_k^i > \overline{y}^i. \end{cases} \tag{4.3}$$

Individual demand for differentiated product j is determined by

$$c(j) = \begin{cases} 0, & p(j) > z_k^i \\ 1, & p(j) \leq z_k^i \end{cases} \tag{4.4}$$

where $z_k^i \equiv 1/\lambda_k^i$ denotes the willingness to pay of a consumer in region i and group k for product j. Note that $c(j) = 0$ for all j if $y \leq \overline{y}$, and $c(j) = 1$ for some j if $y > \overline{y}$. In other words, poor consumers with income below $\overline{y}$ buy only food. Wealthy consumers with income above $\overline{y}$, spend their residual income $(y - \overline{y})$ on differentiated products. They purchase one unit of a differentiated product if the price of that product does not exceed their willingness to pay. With growing income they consume an expanding variety of products instead of increasing consumption of the same products. Figure 4.1 below shows (a) individual demand (4.3) for food and (b) individual demand (4.4) for differentiated product j.

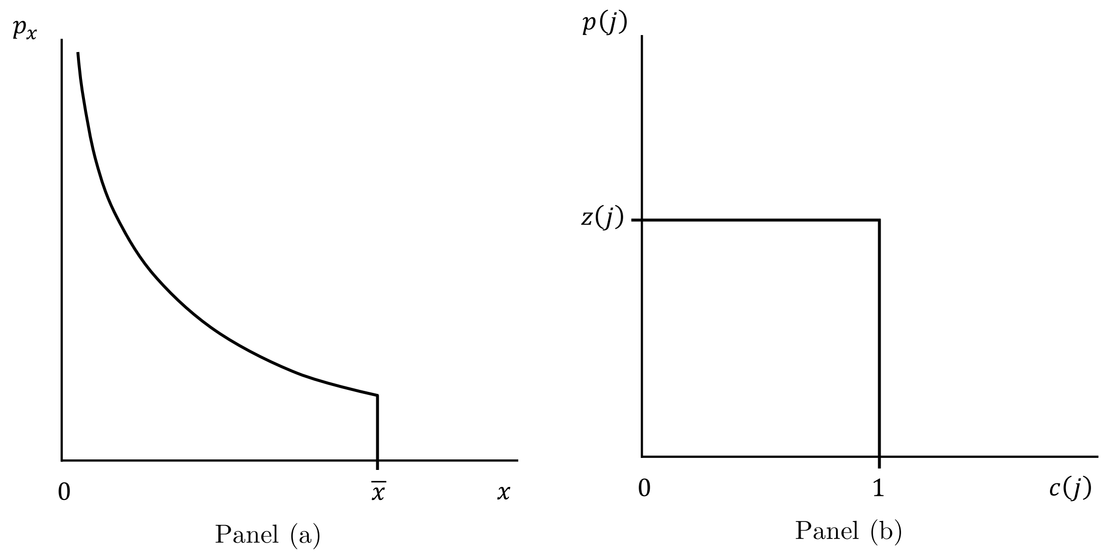


Figure 4.1: Individual demand for (a) food and (b) differentiated product j

4.3.3 Producers

Suppose there is a large number of producers in both regions, which employs (homogenous) labor, the only production factor in the economy.

Homogeneous Good

Food is produced under conditions of perfect competition with the following constant-return-to-scale (CRS) technology. In region i, the production of 1 unit of output requires a^i units of labor. We assume that trade costs in the food sector are prohibitively high, and there is no foreign direct investment.[6]

Differentiated Products

Technology in the differentiated product sector is based on Brainard (1997). Differentiated products are produced under conditions of monopolistic competition with free market entry. Producers have access to the following increasing-returns-to-scale (IRS) technology, which might differ across regions. A producer in region i needs to invest f_E^i units of labor to invent (differentiate) a new product or blueprint. After investing in the creation of a new product a producer in region i must incur f_D^i units of labor to set up a new production plant. The production of 1 unit of output requires b^i units of labor. We assume that producers have a choice between supplying a foreign market by exports or by setting up a foreign production facility, i.e. become a multinational corporation (MNC). By assumption we rule out the possibility of selling or licensing the right to use a blueprint to foreign producers. Suppose that exporters incur all costs in the country of production, whereas multinational producers incur all variable and fixed costs in the foreign production plant in the host country. E.g. a producer with headquarters (HQ) in Home has fixed costs $w^F f_D^H$ to set up a foreign production plant, and variable costs $w^F b^H$ to produce output in the foreign plant. Note that technology is firm-specific and not region-specific. Differentiated products can be traded across regions at iceberg trade costs $\tau \in (1, \infty)$ on each unit shipped. Hence, the model features the familiar proximity-concentration trade-off, where producers would like to maximize the proximity to consumers due to iceberg trade costs, while at the same time, they would like to concentrate their production in one location due to increasing-returns-to-scale technology.

4.3.4 Integrated Equilibrium with FDI and Exports

We are interested in an equilibrium where some producers choose to export and others become multinational corporations. In order to isolate the demand side channel we abstract from heterogeneous technology across regions, i.e. $f_E^i = f_E$, $f_D^i = f_D$, $b^i = b$, and $a^i = a$ for all $i = \{H, F\}$. For an extension discussing differences in technology see Section 4.5.1.

Market Demand

We make the following assumption about the distribution of income across regions and groups.

[6] According to Gibson et al. (2001) the global average tariff (ad valorem) on agricultural products is estimated at 62 percent. See Davis (1998) for a comparison of tariff rates between homogeneous and differentiated goods. FAO (1995) estimates that global agricultural exports account for less than 10 percent of total merchandise exports in 1995. Mahlstein (2010) argues that during the 20th century foreign direct investment in the food sector was negligible.

Assumption 4.2.

$$\left(\theta_1^H - \overline{y}\right) > \tau\left(\theta_1^F - \overline{y}\right) > \tau^2\left(\theta_2^H - \overline{y}\right) > \tau^3\left(\theta_2^F - \overline{y}\right) > 0 > \left(\theta_3^H - \overline{y}\right) > \left(\theta_3^F - \overline{y}\right).$$

This assumption has the following implications. First, the income distribution is such that group 1 corresponds to the rich class, group 2 to the middle class, and group 3 to the poor class in region i. Second, the poor class in both regions cannot afford to purchase differentiated products. Third, income differences across regions and groups are sufficiently high such that producers of differentiated products cannot perfectly price discriminate. For a detailed discussion on the price setting behavior of monopolistic producers see below. Obviously, the equilibrium structure depends crucially on Assumption 4.2. We think that this assumption reflects reality well, and thus constitutes an interesting case worth investigating.

Given Assumption 4.2 aggregate demand is determined as follows. Market demand for food in region i is equal to

$$X^i = \sum_k x_k^i = \overline{x}\left(s_1^i + s_2^i\right)L^i + \frac{y_3^i s_3^i L^i}{p_x^i} \tag{4.5}$$

where the first term denotes aggregate consumption of the middle and rich classes in region i, and the second term total consumption of the poor class. Due to symmetry, market demand is the same for all differentiated products j, and is given by

$$C^i = \begin{cases} 0, & p^i > z_1^H \\ s_1^H L^H, & z_1^F < p^i \leq z_1^H \\ s_1^H L^H + s_1^F L^F, & z_2^H < p^i \leq z_1^F \\ \left(s_1^H + s_2^H\right)L^H + s_1^F L^F, & z_2^F < p^i \leq z_2^H \\ \left(s_1^H + s_2^H\right)L^H + \left(s_1^F + s_2^F\right)L^F, & p^i \leq z_2^F. \end{cases} \tag{4.6}$$

If the price exceeds the willingness to pay of the rich class in Home demand for a differentiated product is zero. If the price is between the willingness to pay of the rich class in Foreign and the rich class in Home demand is equal to the size of the rich class in Home, and so forth. Demand is equal to the size of the middle and rich classes in both regions if the price is equal to or less than the willingness to pay of the middle class in Foreign. Market demand for any product j is depicted in Figure 4.2 below.

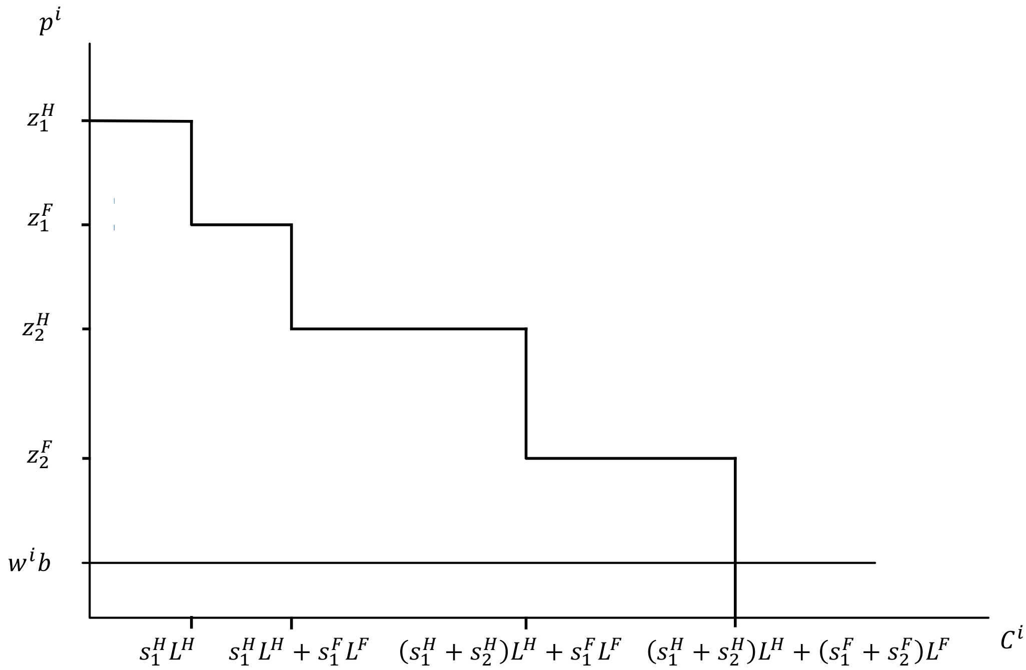


Figure 4.2: Market demand for product j

Aggregate Supply

Perfect competition in the non-traded food sector implies that prices are equal to marginal costs. Hence, the price of the homogeneous good in region i is determined by

$$p_x^i = w^i a. \tag{4.7}$$

Price setting in the differentiated product market is non-trivial. Monopolistic competition implies that producers have price setting power, and charge a markup on marginal costs. In general, producers maximize profits subject to their market demand (4.6) by setting a price where marginal revenue equals marginal cost. Hence, producers would like to sell to all consumers, and perfectly price discriminate by charging prices equal to the willingness to pay of each group of consumers. However, the imminent threat of arbitrage opportunities, i.e. the threat of parallel imports, imposes a price setting restriction on producers. If income differences across groups and regions are sufficiently pronounced the price setting restriction becomes binding.[7] Assumption 4.2 implies that income differences are sufficiently large such that producers' price setting is restricted. In particular, Assumption 4.2 implies that $z_1^H \geq \tau z_1^F$ holds in equilibrium. Hence, producers supplying the rich class in Home and Foreign cannot perfectly price discriminate. For illustration, suppose that producer j would set a price equal to z_1^H in Home and equal to z_1^F in Foreign. However, since by Assumption 4.2, $z_1^H \geq \tau z_1^F$ holds, it is a profitable enterprise for arbitrageurs to purchase product j at a low price z_1^F in Foreign

[7] We assume that producers can never perfectly price discriminate within region i since they cannot distinguish between consumers belonging to different groups or storage of such information is prohibitively expensive.

and ship it to Home at costs $\tau > 1$. There it could be sold with a profit at a price marginally below z_1^H. This threat of parallel imports induces the original producer to set a limit price equal to τz_1^F in Home. Similarly, $z_1^F \geq \tau z_2^H$ implies that producers selling to both middle and rich classes in Home and the rich class in Foreign cannot perfectly price discriminate, and set prices equal to z_2^H in Home and τz_2^H in Foreign. Likewise, $z_2^H \geq \tau z_2^F$ means that producers that sell to both middle and rich classes in Home and Foreign cannot perfectly price discriminate, and charge prices equal to τz_2^F in Home and z_2^F in Foreign. Only producers that serve exclusively the rich class in Home can perfectly price discriminate and set prices corresponding to z_1^H.

Pricing strategy

Assumption 4.2 further implies that *some* producers *must* exclude the middle class in Foreign, some the middle class in Home, and some the rich class in Foreign. For example, consider the case where all producers serve both the middle and rich classes in both regions. In that case, consumers in the middle and rich classes in Home pay prices τz_2^F and consumers in the middle and rich classes in Foreign pay prices z_2^F for all differentiated products. This implies that consumers in the middle and rich classes in Home and consumers in the rich class in Foreign would not exhaust their budget constraints. Since they spend all additional income above $\overline{y}^i$ on differentiated products their (marginal) willingness to pay would become infinitely large. This would induce some producers to deviate and exclude the middle class in Foreign. A similar argument applies for the exclusion of the middle class in Home, and the rich class in Foreign. Hence, ex-ante *identical* producers will choose *different* pricing strategies in equilibrium, serving different segments of the market. Note that if Assumption 4.2 does not apply, some producers can perfectly price discriminate across regions, and might not exclude some groups. In the extreme case, where income differences across regions and groups are sufficiently small all differentiated products might be available on the world market to all consumers with income above the income threshold $\overline{y}^i$. In the following, we stick with Assumption 4.2 since we believe that it reflects a more interesting and realistic situation.

Location

Let us take the organizational decision, i.e. whether a firm sets up a foreign production facility or exports, as given for the moment, and look at the location decision. First, consider a producer who decides to serve the middle and rich classes in both regions. Due to Assumption 4.2 she sets prices equal to τz_2^F in Home and z_2^F in Foreign. Suppose this producer decides to engage in FDI and locate her HQ in Home. In that case, she earns profits equal to

$$\pi_M^H = \left(\tau z_2^F - bw^H\right)\left(s_1^H + s_2^H\right)L^H + \left(z_2^F - bw^F\right)\left(s_1^F + s_2^F\right)L^F - \left[w^H\left(f_E + f_D\right) + w^F f_D\right]$$

where the first term denotes profits (i.e. revenue minus variable costs) from sales in the domestic market, the second term profits from foreign affiliate sales, and the last term are fixed costs. Similarly, profits of a producer who chooses to serve the market in Home through foreign affiliate sales and set up her HQ in Foreign are given by

$$\pi_M^F = \left(z_2^F - bw^F\right)\left(s_1^F + s_2^F\right)L^F + \left(\tau z_2^F - bw^H\right)\left(s_1^H + s_2^H\right)L^H - \left[w^F\left(f_E + f_D\right) + w^H f_D\right].$$

It is straightforward to show that $\pi_M^H > \pi_M^F$ if and only if $w^H < w^F$, and vice versa. In an equilibrium where multinational firms locate in both regions, profits must equalize, i.e. $\pi_M^H = \pi_M^F$. Hence, wages must be equalized across regions, i.e. $w^H = w^F$. Let us choose labor in Home as the numeraire and set w^H equal to one. Second, consider a producer who chooses to sell her product to the middle and rich classes in Home but only to the rich class in Foreign. Assumption 4.2 implies that she offers her product at prices z_2^H in Home and at τz_2^H in Foreign. Suppose this producer decides to locate in Home and export to Foreign. That way she makes profits given by

$$\pi_{X,2}^H = \left(z_2^H - bw^H\right)\left(s_1^H + s_2^H\right)L^H + \left(\tau z_2^H - \tau bw^H\right)s_1^F L^F - w^H\left(f_E + f_D\right).$$

Suppose that such a producer located in Foreign opts to serve the market in Home by engaging in FDI. This producer collects profits equal to

$$\pi_{M,2}^F = \left(\tau z_2^H - bw^F\right)s_1^F L^F + \left(z_2^H - bw^H\right)\left(s_1^H + s_2^H\right)L^H - \left[w^F\left(f_E + f_D\right) + w^H f_D\right].$$

At equalized wage rates normalized to 1 it can be shown that $\pi_{X,2}^H > \pi_{M,2}^F$ if and only if $f_D/s_1^F L^F > (\tau - 1)b$. In that case, all producers who choose this pricing strategy locate in Home. Third, consider a producer who supplies only the rich class in both regions. If Assumption 4.2 holds this producer puts her product up for sale at prices τz_1^F in Home and z_1^F in Foreign. Suppose that this producer locates in Home and exports to Foreign. In doing so, she makes profits given by

$$\pi_{X,1}^H = \left(\tau z_1^F - bw^H\right)s_1^H L^H + \left(z_1^F - \tau bw^H\right)s_1^F L^F - w^H\left(f_E + f_D\right).$$

Similarly, such a producer who decides to locate in Foreign and export to Home obtains profits equal to

$$\pi_{X,1}^F = \left(z_1^F - bw^F\right)s_1^F L^F + \left(\tau z_1^F - \tau bw^F\right)s_1^H L^H - w^F\left(f_E + f_D\right).$$

Having the U.S. in mind as Home and less affluent regions as Foreign, we make the assumption that the size of the middle and rich class, respectively, is larger in Home than in Foreign.

Assumption 4.3. *$s_1^H L^H > s_1^F L^F$, and $s_2^H L^H > s_2^F L^F$.*

If Assumption 4.3 holds, it is straightforward to show that $\pi_{X,1}^H > \pi_{X,1}^F$, with wage rates equalized across regions. Hence, Assumptions 4.2 and 4.3 jointly imply that all producers opting for this pricing strategy locate in Home. Eventually, consider a producer who sells exclusively to the rich class in Home. This producer can perfectly price discriminate, and market her product at a price equal to z_1^H. If she locates in Home she receives profits given by

$$\pi_1^H = \left(z_1^H - bw^H\right) s_1^H L^H - w^H \left(f_E + f_D\right).$$

At equal wage rates across regions one can show that in the presence of iceberg trade costs $\tau > 1$ this producer locates in Home.

Organization

Now, we show that no producer has an incentive to deviate from the organizational form we conjectured above. Given there is a large number of producers the behavior of a single producer has no impact on aggregate variables. Comparing profits from different strategies, it is straightforward to show that no producer has an incentive to deviate from its chosen mode to serve the foreign market if and only if the following proposition holds.

Proposition 4.1. *Given the following condition holds for all $i = \{H, F\}$, all producers serving both the rich and middle classes (i.e. groups 1 and 2) in region i become MNCs whereas all producers supplying only the rich class (i.e. group 1) in region i export:*

$$\frac{f_D}{s_1^i L^i} > (\tau - 1)b > \frac{f_D}{\left(s_1^i + s_2^i\right) L^i}. \tag{4.8}$$

Proof. The proof follows from comparing producers' profits under alternative forms of organization. □

Condition (4.8) has a simple economic intuition. It balances the benefit of engaging in FDI against the cost of that choice. The term $(\tau - 1)b$ denotes the cost reduction per unit sold if a producer engages in FDI instead of exporting her product, due to lower variable costs since no transportation costs incur (opportunity cost of FDI). The term $f_D/\left(s_1^i + s_2^i\right) L^i$, $f_D/s_1^i L^i$, respectively, denotes the cost increase per unit sold of engaging in FDI due to the fixed cost associated with setting up a foreign production facility. We note that, ceteris paribus, producers choose FDI over exports if variable costs are high, or fixed costs to set up a foreign production factory are low relative to the number of units sold (market size in terms of consumers). For a detailed discussion about how condition (4.8) compares to the condition in the case of standard CES preferences see Section 4.5.2 and Appendix 4.A.3.

In particular, Proposition 4.1 implies that in equilibrium no producer has an incentive to set up only "headquarters-services" in Home, i.e. create a new product in Home, and produce all output in Foreign. For example, a situation where producers serving the middle and rich classes in both regions incur f_E in Home, i.e. locate their "headquarters-services" in Home, produce all output only in Foreign (i.e. incur f_D only once in Foreign), and ship it back to Home at trade costs $\tau > 1$ cannot occur in equilibrium. Furthermore, together with Assumption 4.3 no producer serving the middle and rich classes in Home and the rich class in Foreign, or the rich classes in both regions, respectively, has an incentive to locate only her "headquarters-services" in Home, and produce all output in Foreign.

At this point, a brief remark about the assumption of indivisible differentiated products is appropriate. While this assumption makes the model tractable it comes at the cost of shutting down the intensive margin of consumption. To better understand its implications suppose for the moment that consumers had a choice about the extensive and intensive margin. This would be the case with e.g. quadratic preferences over differentiated products of the following form $\int_{\mathcal{J}} \left(sc(j) - \frac{1}{2}c(j)^2 \right) dj$, where $0 < s < \infty$ denotes a local satiation level.[8] In that case, wealthy consumers would not only purchase a larger variety of products but also a larger amount of each product relative to less affluent consumers. That would increase the incentives to engage in FDI relative to exporting for all producers, in particular, also those serving the rich market segment. Hence, the parameter space for which condition (4.8) is fulfilled contracts but does not collapse entirely, such that an equilibrium where some producers engage in FDI and others export is still feasible.

Equilibrium Structure

Given Assumptions 4.1-4.3 and Proposition 4.1 the equilibrium structure looks as follows. Non-homothetic consumer demand segments the market for differentiated products in the sense that it makes ex-ante identical producers choose different pricing strategies to supply different market segments. Differentiated products that are sold to the middle and rich classes in both regions are manufactured in both regions. All products which are sold to all consumers above income threshold $\overline{y}^i$ in region i except the middle class in Foreign are produced in Home.[9] Figure 4.3 below summarizes the equilibrium structure, where the number of MNCs with headquarters in region i are denoted by M^i, the number of exporters located in Home that sells to the middle and rich classes in the domestic market and to the rich class in Foreign by $N^H_{X,2}$, the number of exporters located in Home that serves only the rich class in both regions by $N^H_{X,1}$, and the number of producers located in Home which is not internationally active and sells only to the rich in Home by N^H_1. To close the model, we use the resource and consumer budget constraints. See Appendix 4.A.1 for details and the formal solution to the model.

[8] With $s < \infty$ consumption of an infinitesimal amount of some product j does not yield infinite utility, and therefore generates a non-trivial extensive margin of consumption.

[9] In Appendix 4.A.1 we show that Foreign's trade deficit is accommodated by a surplus in net factor payments.

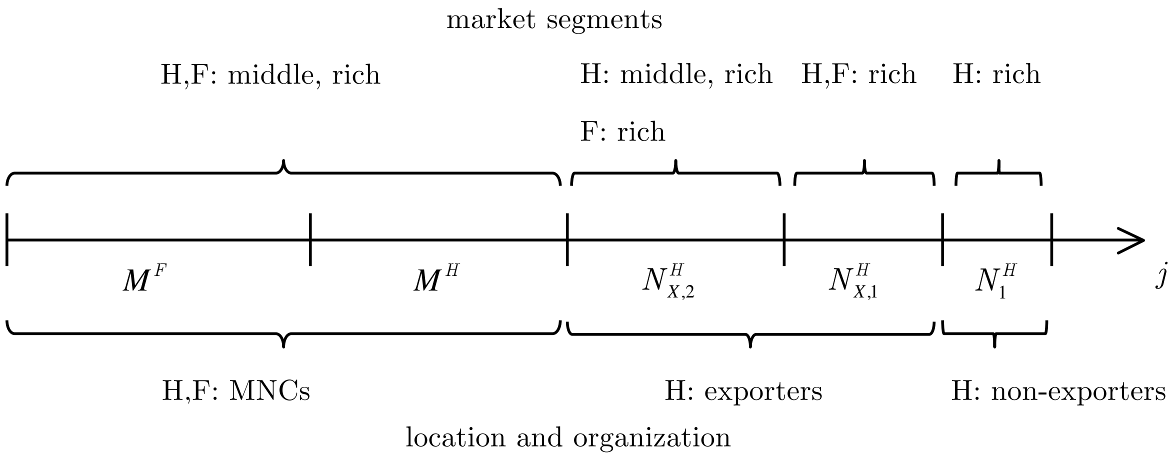


Figure 4.3: Equilibrium structure

4.4 Income Distribution, Market Size, and FDI vs. Exports

We now analyze in detail how market segmentation determined by the income distribution affects foreign direct investment and exports. We will take the viewpoint of Home, and ask how varying market sizes in Foreign, affect the incentive to engage in FDI for producers in Home. In particular, we will focus our analysis on the following two experiments. First, we change the size of the market by changing per capita incomes θ_2^F of consumers in the middle class in Foreign, and second we change the size s_2^F of the middle class in Foreign. In both experiments we hold aggregate income $Y^i \equiv \sum_k s_k^i L^i \theta_k^i$, and average income Y^i/L^i for all i, k constant. We assume that Assumptions 4.2 and 4.3, and Proposition 4.1 holds in all experiments. The intuition for the first case is discussed in some detail, whereas the discussion for the remaining cases is kept short as the intuition is similar.

4.4.1 Changing Per Capita Income of the Middle Class in Foreign

Let us start by considering an increase in per capita income of the middle class θ_2^F, where we (i) decrease per capita income of the rich class θ_1^F, and (ii) decrease per capita income of the poor class θ_3^F, holding everything else constant.

Case (i): Redistribution from Rich to Middle Class

We transfer incomes from the rich class to the middle class. One can show that this increases the number of MNCs with HQ in Home, M^H, whereas the number of MNCs with HQ in Foreign, M^F, decreases. However, the total number of MNCs in the world, $\sum_i M^i$, unambiguously increases. The number of exporters $N_{X,2}^H$ serving the middle and rich classes in Home and only the rich class in Foreign decreases, whereas the number of exporters $N_{X,1}^H$ serving only the rich classes in both regions, and the number of non-exporters N_1^H might increase or decrease.

The intuition is the following. A higher θ_2^F implies that the willingness to pay of consumers belonging to the middle class in Foreign, and therefore prices for products sold in the mass market (i.e. to middle and rich classes in both regions) rises. This creates a disequilibrium in the product market. At higher prices producers who sell in the mass market make positive profits, ceteris paribus. Hence, some producers decide to enter that market segment. Ceteris paribus, this implies that expenditures on mass products increase, depressing expenditures on exported and non-exported products made in Home. This further implies that labor demand in Foreign increases more than labor demand in Home so that relative wages w^H/w^F fall below one. At relative wage rates $w^H/w^F < 1$ profits of MNCs with HQ in Home exceed profits of those with HQ in Foreign, i.e. $\pi_M^H > \pi_M^F$. Hence, producers that enter the mass market set up their HQ in Home. This ameliorates the disequilibrium in the labor market by increasing labor demand in Home. However, at the same time labor demand in Foreign also increases since MNCs absorb resources in the host region. Because labor supply in Foreign has not changed, MNCs with HQ in Foreign start exiting the market. In other words, MNCs with HQ in Home crowd out MNCs with HQ in Foreign. Nevertheless, in equilibrium more MNCs enter in Home than exit in Foreign. Note that exit and entry is such that prices for products sold in the mass market are the same in the old and new equilibrium. In other words, competition intensifies to such a degree that prices fall to the old equilibrium level.

Next, consider the changes induced by the income transfer in the number of exporters and non-exporters. Notice that expenditures on mass products have increased for all consumers above the income threshold $\overline{y}$. First, consider the budget constraint of the middle class in Home, which are the decisive consumers for products sold by producers $N_{X,2}^H$ to middle and rich classes in Home and the rich class in Foreign. Since their income has not changed but their expenditures on mass products has increased they must decrease their expenditures on products sold by producers $N_{X,2}^H$. Hence, some of those producers exit the market, i.e. $N_{X,2}^H$ decreases. Next, consider the budget constraint of a consumer belonging to the rich class in Foreign. She is the decisive consumer for products sold exclusively to the rich classes in both regions. On the one hand, her expenditures for mass products increase by 1, and on the other hand, decrease by τ^2, for products sold to the middle and rich classes in Home and the rich class in Foreign. The net change in her expenditures on those products is negative and equal to $\left(\tau^2 - 1\right)$. This tends to increase her expenditure on products sold exclusively to the rich classes in both regions. At the same time, her income θ_1^F falls by s_2^F/s_1^F, which tends to decrease her expenditures on exclusive products sold only to the rich classes. Hence, if $\left(\tau^2 - 1\right) < s_2^F/s_1^F$ total expenditure on products which are sold only to the rich classes falls, and the number of producers, $N_{X,1}^H$, catering to that market segment decreases in equilibrium. Last, looking at the budget constraint of rich consumers in Home reveals that if their expenditures on products sold exclusively to the rich classes falls, their expenditures on products sold exclusively to them must increase, and vice versa. In that case, the number of producers, N_1^H, serving that market segment rises. We conclude that the effect on the trade volume, i.e. the value of exports from Home to Foreign, given by $\left(N_{X,2}^H \tau z_2^H + N_{X,1}^H z_1^F\right) s_1^F L^F$, is ambiguous. Of course, producers exit and enter the different market segments across the two regions until the equilibrium is restored at relative wage rates $w^H/w^F = 1$.

Case (ii): Redistribution from Poor to Middle Class

We redistribute income from the poor to the middle class. It can be shown that the number of MNCs with HQ in Home, M^H, increases and with HQ in Foreign, M^F, decreases. However, the total number of MNCs $\sum_i M^i$ again increases. At the same time, the number of exporters, $N^H_{X,2}$, that sell to the middle and rich classes in Home and solely to the rich class in Foreign decreases, the number of exporters, $N^H_{X,1}$, catering exclusively to the rich classes in both regions increases, and the number of non-exporters N^H_1 selling only to the rich class in Home falls.

The intuition is similar to case (i) above. Thus, we keep the discussion short and emphasize the difference. First, the intuition for the change in the number of MNCs is the same as before.[10] However, the number of exporters, $N^H_{X,1}$, that sell exclusively to the rich classes in both regions unambiguously increases in equilibrium since the income of the rich class in Foreign does not decrease in case (ii). Thus, because their expenditures on mass products and products sold to the middle and rich classes in Home and the rich class in Foreign falls, they increase their expenditures on products sold by producers $N^H_{X,1}$. This implies that $N^H_{X,1}$ increases. However, the rich class now spends more on all products sold at least to two classes above income threshold $\overline{y}$, and therefore has less income to spend on products sold exclusively to them. Hence, the number of non-exporters, N^H_1, serving that market segment decreases. Again, the effect on the trade volume is ambiguous.

We summarize our main results in the following proposition

Proposition 4.2. *An increase in per capita incomes θ^F_2 of the middle class in Foreign leads to an increase in the number of MNCs with HQ in Home, and has an ambiguous effect on the volume of exports from Home to Foreign.*

Proof. The proof follows from writing $\theta^F_k = \left(\overline{Y} - \theta^F_2 s^F_2 L^F - \theta^F_{l\neq k} s^F_{l\neq k} L^F\right) / s^F_k L^F$ for $k, l = \{1, 3\}$, where we hold aggregate income $\overline{Y}$ constant, and differentiating the solutions for the number of producers in Appendix 4.A.1 with respect to θ^F_2. □

4.4.2 Changing the Size of the Middle Class in Foreign

Next, consider a decrease in the (relative) size s^F_2 of the middle class in Foreign where we (iii) adjust per capita income θ^F_1 and the size s^F_1 of the rich class in Foreign, and (iv) change per capita income θ^F_3 and the size s^F_3 of the poor class such that aggregate income Y^i and class sizes $\left(s^F_1 + s^F_2\right)$, $\left(s^F_2 + s^F_3\right)$, respectively, are constant. In other words, in this experiment we hold per capita income of the middle class in Foreign constant and change the number of people in the middle class by changing the number of (iii) rich consumers and (iv) poor consumers and their corresponding per capita incomes, holding everything else constant.

[10] Note that the change in demand for food from the poor class has no effect on the wage rate within Foreign. The reason is that wage rates across food and differentiated product sector are equalized because labor within Foreign is perfectly mobile, and aggregate labor supply is constant.

Case (iii): Increasing the Size of the Middle Class, and Decreasing the Size of the Rich Class

Let us start with the case where we decrease the size of the rich class in Foreign and increase the size of the middle class, holding their per capita income constant. We want to hold aggregate and average incomes constant, so we have to adjust per capita income of the rich class accordingly. It can be shown that the number of MNCs in both regions, M^H and M^F, do not change, whereas the number of exporters $N^H_{X,2}$ decreases, $N^H_{X,1}$ might increase or decrease, and the number of non-exporters N^H_1 decreases. The intuition is similar to case (i). However, now total market size for multinational producers does not change. The reason is that on the one hand, prices for their products do not change since per capita income of the middle class is constant, and on the other hand, the size of the market in terms of the number of people that consume their product does not change because the number of people in the middle and rich class is constant. Next, consider the market for products sold to the middle and rich class in Home and the rich class in Foreign. Since the size of the rich class in Foreign decreases less exporters enter that market segment in Home, so that $N^H_{X,2}$ decreases.[11] Whether the number of exporters serving only the rich classes in both regions increases or decreases depends on whether the decrease in the size of the market in terms of customers (since s^F_1 decreases) served outweighs the increase in terms of prices (since θ^F_1 increases). However, in equilibrium expenditures of the rich class in Home on products sold exclusively to them unambiguously decrease. This implies that N^H_1 falls.

Case (iv): Increasing the Size of the Middle Class, and Decreasing the Size of the Poor Class

Now, we increase the size of the middle class in Foreign while at the same time we decrease the size of the poor. Again, since we hold aggregate and average incomes constant, we must adjust per capita income of the poor. In that case, one can show that assuming a price elasticity of z^F_2 with respect to the population size s^F_2 greater than one is sufficient for the number of MNCs with HQ in Home to increase, with HQ in Foreign to decrease, and the total number of MNCs that are active in the world to increase. The number of exporters and non-exporters in Home is unaffected.

The intuition behind the result is best understood with the following argument. The number of exporters and non-exporters are not affected because aggregate expenditures on their products do not change. However, aggregate spending on mass products changes because the size $s^F_2 L^F$ of the mass market in terms of customers has increased. This induces producers to enter the mass market. If the price elasticity of products marketed on the mass market with respect to population size s^F_2 exceeds one, labor demand in Foreign increases more than in Home. This implies that relative wages w^H/w^F fall below one, so that MNCs set up their HQ in Home. The rest of the argument is the same as in case (i). However, entry and exit

[11] Notice that a decrease in the market size in terms of the number of customers induces some producers to exit. Due to less intense competition prices rise such that aggregate expenditures are constant.

of MNCs is such that in equilibrium prices z_2^F are lower in the new equilibrium due to more intense competition.

We summarize our main results in the following proposition

Proposition 4.3. *Enlarging the size of the middle class $s_2^F L^F$ by increasing its population share has an ambiguous effect on the number of MNCs with HQ in Home, and the volume of exports from Home to Foreign.*

Proof. The proof follows from writing $\theta_k^F = \left(\overline{Y} - \theta_2^F s_2^F L^F - \theta_{l\neq k}^F s_{l\neq k}^F L^F\right) / s_k^F L^F$ for $k, l = \{1, 3\}$, where we hold aggregate income $\overline{Y}$ and $\left(s_k^F + s_2^F\right)$ constant, and differentiating the solutions for the number of producers in Appendix 4.A.1 with respect to s_2^F. □

4.5 Extensions

The baseline model is very stylized, and thus helps us understand the role of the demand side by isolating it from other factors like e.g. technology. In this section, we discuss extensions of the baseline model with regard to technology and preferences. First, we extend the baseline model towards heterogeneity in technology by arguing that producers in Foreign might not have access to foreign direct investment (e.g. they do not have the managerial know-how to provide headquarter services). The second extension compares the baseline model to the standard model with CES preferences, and highlights the differences. Last, we confront the objection that a factor endowment model, where the skill distribution is reflected in the income distribution, might be observationally equivalent to our baseline model.

4.5.1 Technological Differences: A North-South Perspective

In this section, we look at a simple extension by assuming that firms in Foreign have no access to foreign direct investment. If they want to serve the foreign market they have to do so by exports. That puts the model into a North-South context where we think of Home as North, and Foreign as South.

For comparison's sake, we assume that parameters are such that the equilibrium structure is similar to the one discussed in the baseline model. To preserve space, we refer the reader to Appendix 4.A.2 for the formal model including its solution. The equilibrium which emerges looks as follows. Producers which supply both the middle and rich classes still locate in both regions. However, those located in the South can only serve the foreign market through exports whereas producers in the North set up headquarters there and engage in FDI in the South. Relative wage rates $\omega \equiv w^H/w^F$ are such that producers with that pricing strategy break even, regardless of whether they choose to locate in North or South. In the equilibrium we consider, relative wages ω exceed one. This is the case if labor in the North is more productive than in the South, which we think is a reasonable assumption in this context. In other words, if a North producer's amount of labor used to supply the foreign market, $f_E + f_D + b\left(s_1^H + s_2^H\right) L^H$, falls short of a South producer's labor input, $f_E + \tau b\left(s_1^H + s_2^H\right) L^H$, to supply the market in the North. In sum, $\omega > 1$ if and only if $(\tau - 1)b > f_D / \left(s_1^H + s_2^H\right) L^H$, which is identical to condition (4.8) for i equal to H. In the baseline model, labor productivity is the same for all producers. Therefore, in an equilibrium where MNCs are present in both regions (relative) wages are equalized. Note that relative labor productivity, and therefore the relative wage rate, is determined by the inverse of relative labor requirements given above. Furthermore, we assume that relative wages are such that all producers excluding the middle class in the South choose to locate in the North. This is identical to the equilibrium structure in the baseline model. However, this implies that there is an upper bound ϕ_2 on relative wage rates, where ϕ_2 is defined as the labor productivity of a producer located in the North selling to the rich class in both regions relative to such a producer located in the South. For a formal definition see Appendix 4.A.2. Hence, if $\phi_2 > \omega > 1$ holds, no producer has in equilibrium an incentive to locate in a different region. Similar to the baseline model, no producer in the North has an incentive to deviate from its chosen mode to serve the foreign market if and only if

$$\frac{f_D}{s_1^F L^F} > (\tau\omega - 1)\, b > \frac{f_D}{\left(s_1^F + s_2^F\right) L^F}$$

where ω is endogenously determined by $\left[f_E + \tau b \left(s_1^H + s_2^H\right) L^H\right] / \left[f_E + f_D + b \left(s_1^H + s_2^H\right) L^H\right]$. The condition above is the equivalent to condition (4.8) in the baseline model. By assumption, producers in the South have no choice but to export. In sum, we get a similar equilibrium structure in the North-South context as in the baseline model, except that now there are no Southern MNCs by assumption. Notice that the same equilibrium structure would emerge if we assumed identical technologies across countries but modified condition (4.8) such that all producers located in the South that sell to the middle and rich classes in both regions choose exports over FDI. A Southern producer pursuing that pricing strategy would do so if and only if $f_D / \left(s_1^H + s_2^H\right) L^H > \left(\tau\omega^{-1} - 1\right) b$, which means that the cost of serving the foreign market through exports, i.e. $w^F \tau \left(s_1^H + s_2^H\right) L^H$, is lower than the cost of serving it through a foreign production facility, i.e. $w^H f_D + w^F b \left(s_1^H + s_2^H\right) L^H$. This implies together with the deviation incentive of Northern producers that relative wages ω must be larger than one in equilibrium. We conclude that if labor in the South is sufficiently unproductive relative to labor in the North, engaging in foreign direct investment is less profitable for Southern producers serving the middle and rich classes in both regions compared to exporting.

Again, let us take the viewpoint of producers in the North. The effects of market size on the decision of Northern producers to engage in FDI versus exporting are identical to the baseline model. Thus, we keep the discussion brief and just summarize the results. For the formal treatment refer to Appendix 4.A.2. We start with case (i) and (ii), where we increase per capita income of the middle class in the South and decrease per capita income of the rich, respectively the poor in the South, holding South's aggregate income constant. In both cases, the results and their intuition are exactly the same as in the baseline model. In particular, the number of MNCs with HQ in the North increases in both cases. Remember, in cases (iii) and (iv) we expand the size of the middle class in the South and reduce the size of the rich, respectively, the poor class in the South, while holding aggregate and average income in the South constant. Both, case (iii) and case (iv) are also identical to the baseline model. Hence, we conclude that the results of the baseline model are robust with respect to the type of technological heterogeneity discussed in this section.

4.5.2 Homothetic Consumer Preferences

In this section, we briefly compare the baseline model discussed above to the standard model with homothetic preferences, see e.g. Brainard (1997). We restrict the discussion to the basic intuition and refer to Appendix 4.A.3 for the formal model. Consider homothetic preferences given by $U = x^{\beta}C^{1-\beta}$, where $C = \left(\int_{\mathcal{J}} c(j)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$ with $\sigma > 1$. In other words, consumers have Cobb-Douglas preferences over food and differentiated products, and a constant-elasticity-of-substitution (CES) subutility, which aggregates differentiated products into a composite good C. Cobb-Douglas preferences imply that consumers spend a constant share β of their income on food, and the rest on differentiated products. One can show that with CES subutility all consumers in both regions purchase all differentiated products available, even in the presence of (finite) trade costs. Consumers with different income levels differ only with respect to the amount they buy of each product. In this sense, the CES utility function restricts a consumers choice to the intensive margin of consumption, leaving her no choice about variety. In Appendix 4.A.3, we show that due to the assumption of homothetic preferences the distribution of income within regions has no effect on aggregate demand, and therefore on the number of exporters and MNCs, respectively. Aggregate demand only depends on aggregate income of a region. Furthermore, we argue that a mixed equilibrium where some producers export and others engage in FDI exists only under a knife-edge condition. Hence, the number of exporters and multinationals is indeterminate. In sum, in the standard model it is meaningless to distinguish between different market segments since the income distribution is irrelevant.

4.5.3 Skill vs. Income Distribution

We argued in our baseline model that demand side effects lead to the emergence of MNCs. In equilibrium, the distribution of income affects a producer's decision to serve a particular segment of the (foreign) market either by engaging in FDI or by exporting. We showed in the previous section that such an equilibrium only exists under knife-edge conditions if we shut down the demand channel in the baseline model by assuming homothetic preferences.

However, one could argue that the income distribution reflects the distribution of skills in the economy, i.e. there is a mapping of the skill to the income distribution. Thus, we construct a simple factor proportions model with homothetic consumer behavior where producers combine low, medium, and high-skilled labor in the production of goods, so that (relative) differences in the skill distribution of regions determines the emergence of multinationals versus exporters. Changes in the skill, and therefore income distribution have no effect on aggregate demand but affect aggregate supply by changing (relative) production costs. The question is whether the link between skill, respectively, income distribution is different from our baseline model in a way that could be taken into account if we go to the data. Our model is based on Markusen and

Venables (2000), and Egger and Pfaffermayr (2005). See Appendix 4.A.4 for the formal model. Markusen and Venables (2000) argue in a similar model with skilled and unskilled labor that MNCs are more common when countries are similar in absolute and relative factor endowments. Egger and Pfaffermayr (2005) discuss a model similar to Markusen and Venables (2000) but with three different production factors (skilled and unskilled labor, and physical capital). They show that an increase in the endowment of skilled labor and capital, respectively, of country i relative to country l leads to an increase, and a decrease, respectively, in the number of MNCs relative to exporters in country i. One key result of their analysis is that whereas exports and foreign affiliate sales increase with the similarity in country size, FDI (defined as capital exports) increase with the size of the source country.

Nevertheless, both do not discuss the association between changes in the host country's income distribution induced by changes in its skill distribution, and its effect on the number of multinationals, respectively exporters, active in the parent country. In Appendix 4.A.4 we simulate and discuss in detail the same experiments (i)-(iv) as in the baseline model. To preserve space, we only state the comparative statics results. Our simulations show that in case (i) there is a positive correlation between per capita income of the middle class in Foreign and the number of MNCs with HQ in Home, whereas in case (ii) there is a negative correlation. Remember, our baseline model predicts a positive association between per capita income of the middle class in Foreign and the number of MNCs with HQ in Home in both case (i) and (ii). The simulations also reveal a positive relationship between the size of the middle class in Foreign and the number of MNCs with HQ in Home for both case (iii) and (iv), whereas the baseline model predicted no effect, and also a positive relationship, respectively. We conclude that the predictions of the two models diverge with respect to the relationship between the market size of the middle class in the host country and FDI activity in the host country.

4.6 Empirics

This section illustrates the market size theory with regard to its predictions about FDI activity. In particular, we focus on the market size of the middle class in the host country and its effect on FDI activity in the host country as predicted by cases (i)-(iv), which are summarized in Proposition 4.2 and 4.3. Based on the theory, we would expect a positive relationship between per capita income of the middle class in the host country and FDI activity in the data. Furthermore, the theory suggests either no association between the size of the middle class in the host country and FDI activity, or a positive one.

4.6.1 Data

We use data on outward FDI positions from the OECD (2012) FDI statistics database. The database covers bilateral outward FDI positions, measured in nominal USD, of all OECD countries in 235 countries (including all OECD countries) over the time period of 1985-2011.[12] In 2011, OECD countries held around 80 per cent of the global outward FDI stock (OECD 2013). To approximate FDI positions in real PPP terms we use a GDP deflator from the World Bank (2012), and a PPP conversion factor from Penn World Tables (Heston et al. 2012).

In the model, FDI activity is reflected in the number of MNCs setting up production plants in the foreign country designed to serve local consumers (horizontal FDI), whereas in the data, we proxy FDI activity with outward FDI positions. First, in the data recorded foreign direct investment is defined as obtaining a lasting interest in an entity resident in an economy other than that of the investor. Usually, the lasting interest is defined as obtaining at least 10 per cent of ordinary or voting stock (OECD 1999). Second, it is impossible to distinguish e.g. vertical from horizontal, or greenfield from brownfield investment in the data. Of course, this approximation of FDI activity is far from perfect. However, due to the availability and quality of aggregate data it is still a reasonable choice.

The construction of the market size measures is based on empirical income distributions kindly provided by Bernasconi (2013). She uses data on deciles and quintiles from the World Income Inequality Database UNU-WIDER (2008) and GDP per capita from Heston et al. (2012), to compute empirical income distributions without making parametric assumptions. In particular, Bernasconi (2013) assigns an average income level to each decile (quintile), and redistributes the corresponding area uniformly on an income interval. The resulting densities are then divided into common income intervals on USD $1, 5000, \ldots, 145000, 150000$. We define global income classes by imposing a common lower $\underline{y}$ and upper threshold $\overline{y}$ on each country's income distribution, where $\overline{y} > \underline{y}$. This creates a low income class with per capita incomes $y \leq \underline{y}$, a middle income class with $\overline{y} \geq y > \underline{y}$, and a high income class with $y > \overline{y}$. We measure the size of the different market segments in the host country as aggregate income Y_k and number of people P_k in each income class $k = \{low, middle, high\}$. For a more detailed description see Appendix 4.A.5 and Bernasconi (2013).

Data on control variables are obtained from various sources. Data on regional trade agreements, customs union, and (relative) skill levels are provided by Blonigen and Piger (2011), on schooling by Barro and Lee (2010), on corporate tax and urban concentration by the World Bank Development Indicators (World Bank 2012), on GDP and GDP per capita, trade openness (exports plus imports divided by GDP), and remoteness (distance of country j from all other countries in the world weighted by those other countries' share of world GDP excluding country j) by Heston et al. (2012), on (geodesic) distance by Mayer and Zignago (2001), and on common language and colonial relationship by Helpman et al. (2008).[13]

Due to limited availability of data on the market size measure and control variables, we finally use data on the outward FDI positions of 29 OECD countries in 66 host countries from

[12] The OECD (1999) recommends market value as the conceptual basis for both the valuation of direct investment stocks (i.e. equity and debt instruments) and flows. Note that the inclusion of inter-company debt and of loans from subsidiaries to parent companies may result in some cases in negative values of direct investment stocks. In the sample this concerns less than 2 per cent of observations, which we drop.

[13] For the following variables yearly data is not available, and we average over the time periods data is available: regional trade agreements, customs union, skill levels, schooling, corporate taxes, and urban concentration.

1997 to 2007. The OECD reports whether a FDI position is missing, zero or confidential. Since we use only data on FDI positions reported either positive or zero, we are left with 11,817 observations. The outward FDI position is zero for about 25 percent of all observations in the sample. In Appendix 4.A.5, the list of countries is given in Table A.1, and summary statistics are provided in Tables A.2 and A.3.

4.6.2 Empirical Model

We estimate the following model with Poisson pseudo-maximum likelihood (PPML)

$$\begin{aligned} \log(FDI_{ijt}) &= \alpha + \beta_1 \log(Y_{middle,jt}) + \beta_2 \log(P_{middle,jt}) + \beta_3 \log(Y_{k,jt}) + \beta_4 \log(P_{k,jt}) \\ &+ X_{ijt}\gamma + A_i + A_t + \varepsilon_{ijt}, \quad k \in \{low, high\} \end{aligned} \tag{4.9}$$

where FDI_{ijt} denotes parent country i's FDI position in host country j at time t. In order to include observations where FDI positions are zero, we estimate equation (4.9) with PPML as proposed by Santos Silva and Tenreyro (2006) in the context of gravity equations for bilateral trade flows. The different market segments are captured by aggregate income $Y_{k,jt}$ and number of people $P_{k,jt}$ in income class $k = \{low, middle, high\}$.[14] We run equation (4.9) either excluding the high or low income class, respectively. The coefficient β_1 can be interpreted as the marginal effect of increasing per capita income of the middle class in the host country (by increasing the middle class' total income), simultaneously decreasing per capita income of the rich or poor class, respectively, such that aggregate and average income in the economy are constant. This corresponds to the experiments in cases (i) and (ii) discussed in Section 4.4. The sum of the coefficients β_1 and β_2 can be interpreted as the marginal effect of increasing the size of the middle class in the host country and holding their per capita income constant, while changing size and per capita income of the rich or poor class, respectively, such that aggregate and average income in the economy remain constant.[15] This corresponds to the experiments in cases (iii) and (iv) discussed in Section 4.4.

The vector X includes the following control variables: (log) host real GDP and real GDP per capita, (log) host remoteness, (log) host urban concentration, (log) host trade openness, (log) host corporate tax, (log) distance, (log) squared skill difference (proxy for relative skilled labor endowments), and dummy variables on common official language, regional trade agreement, customs union, colonial relationship, and host region. The inclusion of these control variables

[14] Note that we transform the data by taking logs. The market size as measured by the number of people (measured in thousands) and their respective aggregate income (measured in millions) has zero incidents. To preserve those information we proceed by adding 1 to each observation. Adding 1 to each observation has some tradition in the literature, see e.g. Eichengreen and Irwin (1998), Limao and Venables (2001), Calderon et al. (2007) or Bloom et al. (2011). We experiment adding 0.1, 0.5, 1.5, or 2 and the results are very similar. However, if we add 1 to each observation for the number of people and the mean of the income class to each observation for aggregate income the results are robust for the narrow definitions of the middle class (see below for the definitions of the middle class) but not for the the broader definitions (wrong signs). Similarly, if we add a dummy variable taking the value zero whenever $\{Y_k, P_k\}$ is zero, and 1 otherwise. Since there is no obvious solution to the problem we decide to follow the literature.

[15] We can write $\beta_1 \log(yP) + \beta_2 \log(P)$ where $y = Y/P$ so that $\beta_1 \log(y) + (\beta_1 + \beta_2) \log(P)$.

is guided by Blonigen and Piger (2011). They use Bayesian statistical techniques (i.e. Bayesian Model Averaging) to select from a large set of control variables used in the literature those that are most likely determinants of FDI activity. From their specification using logged FDI stocks in 2000 as a measure of FDI, we include those covariates with an inclusion probability of 50 per cent or higher, excluding parent country covariates which we capture with a parent country fixed effect A_i. That disciplines the empirical exercise since the baseline model does not provide a structural equation. To address the concern that our market size measure might reflect largely skill levels (see discussion Section 4.5.3), we additionally include the percent of population that completed no schooling, primary, secondary, and tertiary education (in logs) in the host country. We also add a host region fixed effect A_r and a year fixed effect A_t.[16] We include a region fixed effect rather than a host country fixed effect because there is little variation in the measure of market size over time due to its construction (see Appendix 4.A.5). In other words, the identification of the effect of $Y_{k,jt}$ and $P_{k,jt}$ comes mainly from its variation across countries.

Definition of the Middle Class

There is no consensus in the literature on how to define a global middle class, and hence on the choice of the income thresholds $\{\underline{y}, \overline{y}\}$. For a detailed discussion on the definition of the middle class see e.g. Ravallion (2010). We apply three different definitions commonly used in the literature. First, we choose the thresholds according to Milanovic and Yitzhaki (2002), see also World Bank (2007), which define the middle class as those individuals with per capita incomes between USD 4,500 and USD 19,000 (PPP, 2005). The lower bound corresponds to Brazil's per capita income, and the upper bound to Italy's per capita income. This is one of the most widely used definitions of the middle class in the literature. Second, we construct a world income distribution from Bernasconi's (2013) country income distributions, which has median income USD 5,000 and mean income ranging from USD 7,500 to USD 10,100 over the period of 1997-2007. Defining the middle class as mean income $\pm 25\%$ yields a lower threshold of USD 5,000 and an upper threshold of USD 10,000. We use the mean rather than the median since by construction we can vary income thresholds only in USD 5,000 steps. This definition has some tradition in the literature, see e.g. Thurow (1987). Although arbitrary these thresholds have some economic rationale. For example, Muhammad et al. (2011) estimate that households below a per capita income of approximately USD 6,500 (PPP, 2005), which corresponds to a real per capita income less than 15 per cent of the US level, spend on average between 50 and 60 per cent of their income on food and housing. The McKinsey Global Institute (2010) classifies households in the income bracket of USD 10,000 - 20,000 (PPP, 2005) as the consuming middle class, and those with incomes below USD 5,000 as satisfying basic consumer needs. Ali and Dadush (2012) propose the number of passenger cars in circulation as a proxy for the number of people in the middle class. They argue that the car ownership rate accelerates around the per

[16]The World Bank classification uses regions East Asia & Pacific, Europe & Central Asia, Latin America & Caribbean, Middle East & North Africa, South Asia, Sub-Saharan Africa, and Western Europe. We complement Northern America, and Oceania.

capita income threshold USD 3,400 and decelerates sharply once per capita income exceeds USD 25,000 (PPP, 2005). Ali and Dadush (2012) show that their ranking of the middle class across countries is broadly in line with the ranking based on Milanovic and Yitzhaki's (2002) definition. Third, we estimate equation (4.9) for a set of possible combinations of income thresholds, i.e. $\underline{y} \in \{5000, 10000\}$ and $\overline{y} \in \{10000, \ldots, 35000\}$, which includes the two definitions above. In principle, we can then compare the R-squared across different definitions of the middle class.[17] We choose the poverty threshold in the US defined by the US Census Bureau as an upper bound on the lower threshold $\underline{y}$. The United States Census Bureau (2013) considers individuals living in the United States to be poor if their annual income is less than USD 10,160 (in 2005 USD). As an upper bound on the income threshold $\overline{y}$ we choose the median GDP per capita across the 29 OECD countries in the sample, which is equal to USD 34,133 in 2007. According to Ali and Dadush (2012) this is one of the narrowest definitions used in the literature.

4.6.3 Results and Discussion

This section discusses the link between per capita income and the size of the middle class in the host country, respectively, and FDI positions in the host country held by OECD countries.

Changing Income and Size of the Middle Class relative to the Rich Class

Table 4.1 below shows the results from estimating equation (4.9) with PPML, excluding the rich class, and setting the lower threshold at (a) USD 5,000 or (b) USD 10,000, and varying the upper threshold from USD 10,000 to 35,000.

To illustrate the results, consider column one in panel (a) of Table 4.1. First, we find that a 1 per cent increase in per capita income of the middle class in the host country, while simultaneously decreasing per capita income of the rich class, increases the FDI position in the host country held by OECD countries by 2.478 per cent (β_1) on average, holding in particular income and size of the poor class, GDP per capita and GDP constant. Second, we find that a 1 per cent increase in the size of the middle class in the host country, while simultaneously decreasing the size of the rich class, decreases the FDI position in the host country held by OECD countries by 0.575 per cent ($\beta_1 + \beta_2$) on average, again holding income and size of the poor class, GDP per capita and GDP constant. We see that GDP per capita has no effect whereas GDP has a significantly positive effect on outward FDI positions of OECD countries.

These results relate to cases (i) and (iii) discussed in Section 4.4 as follows. They lend support to the prediction of case (i) that an increase in the per capita income of the middle class in the host country, by decreasing per capita income of the rich class, leads to an increase in FDI activity in the host country, which is measured here as the FDI position in the host

[17] We compute a pseudo R^2 as the square of the correlation between the model's predicted values and the actual values, which ranges from 0 to 1 (see Wooldridge 2009). The greater the correlation between the predicted values and the actual values, the greater the R^2. We adjust the R^2 for the number of regressors as follows $1 - \left(1 - R^2\right)(n-1)/(n-p-1)$, where n denotes sample size and p the number of regressors. Even though the pseudo R^2 cannot be compared across datasets it is still valid and useful in evaluating multiple models predicting the same outcome on the same dataset (see Introduction to SAS: UCLA: Statistical Consulting Group 2013).

Table 4.1: Outward FDI positions (29 OECD countries 1997-2007, pooled, PPML)

(a) with lower threshold $\underline{y} = 5,000$

	log(FDI_{ijt})					
	$\overline{y} = 10,000$	$\overline{y} = 15,000$	$\overline{y} = 20,000$	$\overline{y} = 25,000$	$\overline{y} = 30,000$	$\overline{y} = 35,000$
log($Y_{middle,jt}$)	2.478**	1.997**	2.930***	3.300***	3.362***	2.555***
	(1.120)	(0.980)	(0.856)	(0.828)	(0.955)	(0.792)
log($P_{middle,jt}$)	-3.053**	-2.704**	-3.790***	-4.149***	-4.098***	-3.209***
	(1.316)	(1.164)	(1.036)	(1.004)	(1.074)	(0.940)
log($Y_{low,jt}$)	-0.303	-0.256	-0.160	-0.094	-0.121	-0.229
	(0.245)	(0.230)	(0.205)	(0.204)	(0.207)	(0.222)
log($P_{low,jt}$)	0.585	0.522	0.574	0.557	0.511	0.375
	(0.385)	(0.361)	(0.359)	(0.350)	(0.345)	(0.347)
log(host GDP pc$_{jt}$)	0.850	0.669	0.737	0.815	0.824	0.552
	(0.901)	(0.832)	(0.796)	(0.759)	(0.732)	(0.775)
log(host GDP$_{jt}$)	1.066**	1.261***	1.282***	1.215***	1.160***	1.338***
	(0.443)	(0.444)	(0.435)	(0.418)	(0.413)	(0.438)
adj pseudo R^2	0.794	0.802	0.811	0.815	0.816	0.809
#observations	11817	11817	11817	11817	11817	11817

(b) with lower threshold $\underline{y} = 10,000$

	log(FDI_{ijt})					
	$\overline{y} = 10,000$	$\overline{y} = 15,000$	$\overline{y} = 20,000$	$\overline{y} = 25,000$	$\overline{y} = 30,000$	$\overline{y} = 35,000$
log($Y_{middle,jt}$)		2.055***	2.202***	2.547**	3.152	1.694**
		(0.648)	(0.696)	(1.144)	(3.915)	(0.724)
log($P_{middle,jt}$)		-2.702***	-2.862***	-3.154**	-3.482	-2.125**
		(0.802)	(0.874)	(1.246)	(3.326)	(0.893)
log($Y_{low,jt}$)		-0.042	-0.130	-0.223	-0.468	-0.584
		(0.505)	(0.526)	(0.541)	(0.526)	(0.564)
log($P_{low,jt}$)		-0.258	-0.231	-0.142	0.276	0.224
		(0.815)	(0.877)	(0.903)	(1.097)	(0.946)
log(host GDP pc$_{jt}$)		0.159	0.077	0.150	0.672	0.398
		(0.940)	(1.020)	(1.050)	(1.624)	(1.141)
log(host GDP$_{jt}$)		1.752***	1.830***	1.769***	1.299	1.589**
		(0.535)	(0.593)	(0.623)	(1.395)	(0.645)
adj pseudo R^2		0.818	0.816	0.814	0.815	0.806
#observations		11817	11817	11817	11817	11817

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered (by host country) standard errors in parentheses. Excluding rich class. Controls: host remoteness, host urban concentration, host trade openness, host corporate tax, distance, squared skill difference, percent of population that completed no, primary, secondary, and tertiary schooling, dummies for common language, regional trade agreement, customs union, colonial relationship, parent country, host region, and year.

country held by OECD countries. According to case (iii) there should be no effect on FDI activity if we increase the size of the middle class, and hold their per capita income constant (adjusting income and size of the rich class such that aggregate and average income in the economy is constant). However, in the data we see a negative relationship between the size of the middle class in the host country and its FDI positions held by OECD countries. We think of the following possible explanation, which can be reconciled with our baseline model. Suppose for the moment that consumers adjust their consumption not only along the extensive but also the intensive margin. As the number of consumers in the middle class expands more producers enter that market segment, ceteris paribus. However, since per capita income of the middle class has not changed they reduce the units bought of each product in order to afford the additional varieties offered in the market, ceteris paribus. It might be that the negative effects on FDI (less units sold by each producer) outweigh the positive ones (more producers active in the market). Nevertheless, in the data the magnitude of the effect of per capita income of the middle class on the FDI position is about five times larger than the effect of its size.

Table 4.1 shows that these results are robust with respect to the definition of the middle class. The results are similar with respect to estimating equation (4.9) for every year separately. See Table A.4 in Appendix 4.A.5 for details.[18] Of the control variables included in the main specification shown in Table 4.1, remoteness, openness, common language and colonial ties have a significantly positive effect, whereas distance, percentage in the population with completed primary or secondary schooling have a significantly negative effect on outward FDI positions.[19] The signs of the control variables are the same as in Blonigen and Piger (2011).

Changing Income and Size of the Middle Class relative to the Poor Class

Table 4.2 below shows the results from estimating equation (4.9) with PPML, excluding the poor class, and setting the lower threshold at (a) USD 5,000 or (b) USD 10,000, and varying the upper threshold from USD 10,000 to 35,000.

The results in Table 4.2 correspond to cases (ii) and (iv) discussed in Section 4.4. The interpretation is similar to the one in Table 4.1. We see that increasing per capita income of the middle class in the host country and simultaneously decreasing per capita income of the poor, while holding aggregate and average income of the rich class, GDP per capita and GDP constant, leads on average to an increase in the FDI position in the host country held by OECD countries. The magnitude of the increase is between 2.072 per cent and 3.541 per cent for a 1 per cent increase in per capita income of the middle class, depending on the definition of the middle class. This is in line with the prediction of case (ii) in the baseline model. Again, we observe that an increase in the size of the middle class in the host country, holding their per capita income constant, and adjusting income and size of the poor class such that GDP per capita and GDP are constant, implies a decrease in the FDI position of the host country. On average, a 1 per cent increase in the size of the middle class implies a decrease in the FDI

[18] We also estimate (4.9) including a parent country-year fixed effect A_{it} and the results are very similar with respect to the size and significance levels of the coefficients. This is also the case if we exclude the poor class.

[19] Suppose we interpret primary and secondary schooling as medium-skilled labor. In the factor endowment model discussed in Section 4.5.3 increasing the supply of medium-skilled labor decreases its relative price and makes engaging in FDI relative to exporting cheaper, ceteris paribus. Thus, the number of producers engaging in FDI should increase.

Table 4.2: Outward FDI positions (29 OECD countries 1997-2007, pooled, PPML)

(a) with lower threshold $\underline{y} = 5,000$

	log(FDI_{ijt})					
	$\overline{y} = 10,000$	$\overline{y} = 15,000$	$\overline{y} = 20,000$	$\overline{y} = 25,000$	$\overline{y} = 30,000$	$\overline{y} = 35,000$
log($Y_{middle,jt}$)	2.072	-1.322	2.394***	3.355***	3.488***	2.720***
	(1.311)	(1.488)	(0.831)	(0.798)	(0.865)	(0.826)
log($P_{middle,jt}$)	-2.559*	0.569	-3.326***	-4.339***	-4.409***	-3.451***
	(1.479)	(1.544)	(1.007)	(0.983)	(1.050)	(1.009)
log($Y_{high,jt}$)	1.165**	0.482	0.315	0.718	1.012***	0.384*
	(0.582)	(0.572)	(0.862)	(0.538)	(0.352)	(0.210)
log($P_{high,jt}$)	-1.598**	0.495	-0.197	-0.978	-1.593***	-0.742**
	(0.791)	(1.059)	(1.336)	(0.795)	(0.505)	(0.319)
log(host GDP pc$_{jt}$)	-0.017	0.351	-0.201	-0.480	-0.280	0.433
	(0.358)	(0.347)	(0.375)	(0.311)	(0.320)	(0.362)
log(host GDP$_{jt}$)	1.689***	0.579	1.656***	2.078***	2.292***	1.886***
	(0.325)	(0.606)	(0.553)	(0.341)	(0.331)	(0.276)
adj pseudo R^2	0.797	0.810	0.810	0.810	0.821	0.816
#observations	11817	11817	11817	11817	11817	11817

(b) with lower threshold $\underline{y} = 10,000$

	log(FDI_{ijt})					
	$\overline{y} = 10,000$	$\overline{y} = 15,000$	$\overline{y} = 20,000$	$\overline{y} = 25,000$	$\overline{y} = 30,000$	$\overline{y} = 35,000$
log($Y_{middle,jt}$)		2.578***	3.019***	3.541**	3.235**	2.324***
		(0.938)	(0.743)	(1.709)	(1.408)	(0.701)
log($P_{middle,jt}$)		-3.351***	-3.732***	-4.251**	-3.971***	-3.054***
		(1.076)	(0.884)	(1.670)	(1.320)	(0.895)
log($Y_{high,jt}$)		-0.913	-0.771	0.209	0.623**	0.257
		(0.646)	(0.570)	(0.533)	(0.296)	(0.215)
log($P_{high,jt}$)		1.253	0.869	-0.546	-1.131***	-0.611*
		(0.838)	(0.850)	(0.787)	(0.429)	(0.330)
log(host GDP pc$_{jt}$)		0.875***	0.871**	0.592	0.700*	1.063**
		(0.339)	(0.346)	(0.362)	(0.378)	(0.417)
log(host GDP$_{jt}$)		1.255***	1.439***	1.851***	2.018***	1.863***
		(0.226)	(0.324)	(0.364)	(0.349)	(0.284)
adj pseudo R^2		0.821	0.819	0.815	0.824	0.818
#observations		11817	11817	11817	11817	11817

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered (by host country) standard errors in parentheses. Excluding poor class. Controls: host remoteness, host urban concentration, host trade openness, host corporate tax, distance, squared skill difference, percent of population that completed no, primary, secondary, and tertiary schooling, dummies for common language, regional trade agreement, customs union, colonial relationship, parent country, host region, and year.

position between 0.532 per cent and 0.984 per cent. Again, this contradicts the prediction of case (iv), which suggests a positive sign.

From Table 4.2 we see that these results are fairly robust with respect to the definition of the middle class. Only for the definition of the middle class with per capita incomes between USD 5,000 and USD 15,000 are the signs on the market size measures reversed, but not significant. Again, the results are similar with respect to estimating equation (4.9) for every year separately. For details, see Table A.5 in Appendix 4.A.5. The signs and significance levels of the control variables are the same as in the case where we exclude the rich group.

4.6.4 Summary

In sum, we find a positive relationship between per capita income of the middle class in the host country and its FDI stock held by OECD countries, changing per capita income of the rich, respectively, poor class such that aggregate and average income in the economy are constant. Furthermore, we find a negative link between the size of the middle class in the host country and its FDI stock, holding per capita income of the middle class constant, while adjusting the size and income of the rich, respectively, poor class such that GDP and GDP per capita remain constant. This supports the predictions of the baseline model with respect to the effect of the middle class' per capita income in the host country but not with respect to the effect of the size of the middle class. However, we observe that per capita income of the middle class in the host country is a much more important determinant of FDI positions in the host country than the size of the middle class.[20]

We interpret these results as suggestive rather than conclusive evidence for the predictions about FDI activity of the baseline model. Obviously, the purpose of the empirical exercise is to illustrate the baseline model, and cannot be interpreted as a rigorous test of the model. Nevertheless, it is a first encouraging step towards empirically studying demand side effects in the context of FDI determinants. A more serious attempt to test the theory would take the firm's decision to export into account. However, trade flows are clearly endogenous since firms decide simultaneously whether to export or engage in FDI. A solution to this problem would require a valid instrument, with no obvious candidate available. The estimation of the baseline specification (4.9) follows the literature, which has largely ignored the issue (see Blonigen 2005). Additional support for the theory might be found in firm-level data on export and foreign affiliate's sales by destination and host market, respectively. We think this could be a promising avenue worth exploring in the future.

4.7 Conclusion

This chapter investigates how the distribution of aggregate income in the host country affects capital flows/stocks and international trade. Motivated by anecdotal evidence it provides a simple general-equilibrium model where firms choose the foreign market entry mode depending on the market segment they supply. Firms facing a proximity-concentration trade-off choose

[20]Our specification generates an adjusted pseudo R^2 between 0.794 and 0.824 depending on the definition of the middle class. Estimating equation (4.9) excluding the measures for market size yields a coefficient of 0.683 (0.424) on GDP per capita, and of 0.813 (0.134) on GDP, with p-values in parentheses. The adjusted pseudo R^2 is 0.773.

to engage in FDI if they serve the mass market abroad, and decide to export if they sell exclusively to a few rich abroad. For firms serving a large market abroad economies of scale are large enough to compensate for higher fixed costs associated with FDI so that average costs are lower compared to exporting. The analysis highlights the importance of the middle class in the host country in terms of per capita income and size as a determinant of FDI activity in the host country. A point that has been neglected in the previous literature, which attributed no role to the distribution of GDP and approximated market size by GDP or GDP per capita. Using data on outward FDI positions of OECD countries we illustrate the role of per capita income of the middle class in the host country as a determinant of FDI positions in the host country.

4.A Appendix

4.A.1 Baseline Model

This section presents the formal baseline model and its solution.

Resource and Budget Constraints, and the Balance of Payments

The resource constraint in Home, i.e. labor market clearing, is given by

$$\begin{aligned}
\left(\theta_1^H s_1^H + \theta_2^H s_2^H + \theta_3^H s_3^H\right) L^H &= \left[\overline{x}\left(s_1^H + s_2^H\right) L^H + x_3^H s_3^H L^H\right] a \\
&+ M^H \left[f_E + f_D + b\left(s_1^H + s_2^H\right) L^H\right] \\
&+ N_{X,2}^H \left[f_E + f_D + b\left(s_1^H + s_2^H\right) L^H + \tau b s_1^F L^F\right] \\
&+ N_{X,1}^H \left[f_E + f_D + b s_1^H L^H + \tau b s_1^F L^F\right] \\
&+ N_1^H \left[f_E + f_D + b s_1^H L^H\right] + M^F \left[f_D + b\left(s_1^H + s_2^H\right) L^H\right].
\end{aligned}$$

Similarly, the resource constraint in Foreign is given by

$$\begin{aligned}
\left(\theta_1^F s_1^F + \theta_2^F s_2^F + \theta_3^F s_3^F\right) L^F &= \left[\overline{x}\left(s_1^F + s_2^F\right) L^F + x_3^F s_3^F L^F\right] a \\
&+ M^F \left[f_E + f_D + b\left(s_1^F + s_2^F\right) L^F\right] \\
&+ M^H \left[f_D + b\left(s_1^F + s_2^F\right) L^F\right].
\end{aligned}$$

Budget Constraints

Consumers' budget constraints in Home are given by

$$\begin{aligned}
\theta_1^H w^H + v_1^H &= p_x^H \overline{x} + \left(M^H + M^F\right) \tau z_2^F + N_{X,2}^H z_2^H + N_{X,1}^H \tau z_1^F + N_1^H z_1^H \\
\theta_2^H w^H + v_2^H &= p_x^H \overline{x} + \left(M^H + M^F\right) \tau z_2^F + N_{X,2}^H z_2^H \\
\theta_3^H w^H + v_3^H &= p_x^H x_3^H.
\end{aligned}$$

The budget constraints of consumers in Foreign are given by

$$\begin{aligned}
\theta_1^F w^F + v_1^F &= p_x^F \overline{x} + \left(M^H + M^F\right) z_2^F + N_{X,2}^H \tau z_2^H + N_{X,1}^H z_1^F \\
\theta_2^F w^F + v_2^F &= p_x^F \overline{x} + \left(M^H + M^F\right) z_2^F \\
\theta_3^F w^F + v_3^F &= p_x^F x_3^F.
\end{aligned}$$

Note that $w^H = w^F = 1$ and $v_k^i = 0$ for all i, k in equilibrium.

Balance of Payments

The balance of payments is implied by the resource constraint, the zero-profit conditions of producers and the budget constraints of consumers. The balance of payments from the viewpoint of Foreign is given by

$$
\begin{aligned}
- & \left(N_{X,2}^{H}\tau z_{2}^{H}+N_{X,1}^{H}z_{1}^{F}\right)s_{1}^{F}L^{F} \\
+ & \left[\left(\tau z_{2}^{F}-b\right)\left(s_{1}^{H}+s_{2}^{H}\right)L^{H}-f_{D}\right]M^{F}-\left[\left(z_{2}^{F}-b\right)\left(s_{1}^{F}+s_{2}^{F}\right)L^{F}-f_{D}\right]M^{H} \\
= & \ 0
\end{aligned}
$$

where the first two lines denote the current account, which consists of the trade balance (first line) and net factor payments (second line). Note that a static framework cannot account for changes in net foreign asset holdings (due to free entry asset values are zero in a static equilibrium). Thus, the balance of payments corresponds to the current account. The trade balance reports that Foreign's net exports are equal to $\left(N_{X,2}^{H}\tau z_{2}^{H}+N_{X,1}^{H}z_{1}^{F}\right)s_{1}^{F}L^{F}$, which are negative in the case we consider. This trade deficit is accommodated by positive net factor payments, which consist of (total) repatriated profits of MNCs with HQ in Foreign earned from foreign affiliate sales in Home equal to $\left[\left(\tau z_{2}^{F}-b\right)\left(s_{1}^{H}+s_{2}^{H}\right)L^{H}-f_{D}\right]M^{F}$, minus (total) repatriated profits of MNCs with HQ in Home earned from foreign affiliate sales in Foreign equal to $\left[\left(z_{2}^{F}-b\right)\left(s_{1}^{F}+s_{2}^{F}\right)L^{F}-f_{D}\right]M^{H}$.

Solution

We solve for the integrated mixed equilibrium using the budget constraints of consumers and the resource constraint of Foreign. It is easy to check that the resource constraint of Home is redundant due to Walras' law.

Prices

Since technology in the food sector is identical and wage rates equalize across regions the price of food is the same in both regions, and is given by

$$p_{x}^{H}=p_{x}^{F}=a.$$

This implies that the income threshold is the same in both countries and equal to $\overline{y}=a\overline{x}$. From the zero-profit conditions we solve for the willingness to pay of consumers, which are given by

$$
\begin{aligned}
z_{1}^{H} &= \frac{f_{E}+f_{D}+bs_{1}^{H}L^{H}}{s_{1}^{H}L^{H}} \\
z_{2}^{H} &= \frac{f_{E}+f_{D}+b\left(s_{1}^{H}+s_{2}^{H}\right)L^{H}+\tau bs_{1}^{F}L^{F}}{\left(s_{1}^{H}+s_{2}^{H}\right)L^{H}+\tau s_{1}^{F}L^{F}} \\
z_{1}^{F} &= \frac{f_{E}+f_{D}+bs_{1}^{H}L^{H}+\tau bs_{1}^{F}L^{F}}{\tau s_{1}^{H}L^{H}+s_{1}^{F}L^{F}} \\
z_{2}^{F} &= \frac{f_{E}+2f_{D}+b\left(s_{1}^{H}+s_{2}^{H}\right)L^{H}+b\left(s_{1}^{F}+s_{2}^{F}\right)L^{F}}{\tau\left(s_{1}^{H}+s_{2}^{H}\right)L^{H}+\left(s_{1}^{F}+s_{2}^{F}\right)L^{F}}.
\end{aligned}
$$

Assumption 4.2 implies that parameters are restricted such that $z_1^H \geq \tau z_1^F \geq \tau^2 z_2^H \geq \tau^3 z_2^F$ holds. Note that $z_2^F > b$ implies that $(f_E + 2f_D) / \left(s_1^H + s_2^H\right) L^H > (\tau - 1) b$, and $\partial z_2^F / \partial s_2^F < 0$, which is consistent with Proposition 4.1.

Number of Producers

The number of multinational producers who supply all groups above $\overline{y}$ in both regions with headquarters in Home is determined by

$$\begin{aligned} M^H &= \frac{\left(\theta_2^F - \overline{y}\right)\left[\tau\left(s_1^H + s_2^H\right)L^H + \left(s_1^F + s_2^F\right)L^F\right]\left[f_E + f_D + b\left(s_1^F + s_2^F\right)L^F\right]}{\left[f_E + 2f_D + b\left(s_1^H + s_2^H\right)L^H + b\left(s_1^F + s_2^F\right)L^F\right]f_E} \\ &- \frac{\sum_{k=1}^2\left(\theta_k^F - \overline{y}\right)s_k^F L^F}{f_E}. \end{aligned}$$

It can be shown that M^H is increasing in θ_2^F if and only if $\left[f_E + f_D + b\left(s_1^F + s_2^F\right)L^F\right]/z_2^F s_2^F L^F$ exceeds 1, and increasing in s_2^F if the following sufficient (but not necessary) condition holds

$$\varepsilon_{zs}\left[\frac{f_E + f_D + b\left(s_1^F + s_2^F\right)L^F}{z_2^F s_2^F L^F}\right] > \frac{\left(z_2^F - b\right)}{z_2^F}$$

where we assume that the price elasticity $\varepsilon_{zs} \equiv -\left(\partial z_2^F / z_2^F\right) / \left(\partial s_2^F / s_2^F\right)$ with respect to population size is larger than one.

Similarly, the number of producers engaged in FDI with headquarters in Foreign is given by

$$\begin{aligned} M^F &= \frac{\sum_{k=1}^2\left(\theta_k^F - \overline{y}\right)s_k^F L^F}{f_E} \\ &- \frac{\left(\theta_2^F - \overline{y}\right)\left[\tau\left(s_1^H + s_2^H\right)L^H + \left(s_1^F + s_2^F\right)L^F\right]\left[f_D + b\left(s_1^F + s_2^F\right)L^F\right]}{\left[f_E + 2f_D + b\left(s_1^H + s_2^H\right)L^H + b\left(s_1^F + s_2^F\right)L^F\right]f_E} \end{aligned}$$

which is decreasing in θ_2^F if and only if $\left[f_D + b\left(s_1^F + s_2^F\right)L^F\right]/z_2^F s_2^F L^F > 1$, and decreasing in s_2^F if the following sufficient (but not necessary) condition holds

$$\varepsilon_{zs}\left[\frac{f_D + b\left(s_1^F + s_2^F\right)L^F}{z_2^F s_2^F L^F}\right] > \frac{\left(z_2^F - b\right)}{z_2^F}$$

with $\varepsilon_{zs} > 1$.

The total number of producers engaged in FDI across both regions is determined by

$$\sum_i M^i = \frac{\left(\theta_2^F - \overline{y}\right)}{z_2^F} = \frac{\left(\theta_2^F - \overline{y}\right)\left[\tau\left(s_1^H + s_2^H\right)L^H + \left(s_1^F + s_2^F\right)L^F\right]}{f_E + 2f_D + b\left(s_1^H + s_2^H\right)L^H + b\left(s_1^F + s_2^F\right)L^F}$$

which is increasing in θ_2^F, and increasing in s_2^F if and only if $\partial z_2^F / \partial s_2^F < 0$, and vice versa.

Note that the following parameter restrictions apply in an equilibrium where $M^i > 0$ for all i

$$\frac{f_E + f_D + b\left(s_1^F + s_2^F\right)L^F}{z_2^F s_2^F L^F} > \frac{\left(\theta_1^F - \overline{y}\right)s_1^F L^F + \left(\theta_2^F - \overline{y}\right)s_2^F L^F}{\left(\theta_2^F - \overline{y}\right)s_2^F L^F} > \frac{f_D + b\left(s_1^F + s_2^F\right)L^F}{z_2^F s_2^F L^F} > 1$$

where the last inequality holds since repatriated profits of MNCs with headquarters in Home from foreign affiliate sales $\left[\left(z_2^F - b\right)\left(s_1^F + s_2^F\right)L^F - f_D\right]$ are positive (this can also be seen from the resource constraint of Home).

The number of exporters that supply the middle and rich classes in Home and only the rich class in Foreign is given by

$$N_{X,2}^H = \frac{\left[\left(\theta_2^H - \overline{y}\right) - \tau\left(\theta_2^F - \overline{y}\right)\right]\left[\left(s_1^H + s_2^H\right)L^H + \tau s_1^F L^F\right]}{f_E + f_D + b\left(s_1^H + s_2^H\right)L^H + \tau b s_1^F L^F}$$

which is decreasing in θ_2^F, and independent of s_2^F.

Next, the number of exporters which sells only to the rich class in both regions is given by

$$N_{X,1}^H = \frac{\left\{\left[\left(\theta_1^F - \overline{y}\right) - \tau\left(\theta_2^H - \overline{y}\right)\right] + \left(\tau^2 - 1\right)\left(\theta_2^F - \overline{y}\right)\right\}\left(\tau s_1^H L^H + s_1^F L^F\right)}{f_E + f_D + b s_1^H L^H + \tau b s_1^F L^F}$$

which is increasing in θ_2^F, and independent of s_2^F.

Last, the number of those non-exporters that serve only the rich class in Home is determined by

$$N_1^H = \frac{\left\{\left[\left(\theta_1^H - \overline{y}\right) - \tau\left(\theta_1^F - \overline{y}\right)\right] + \left(\tau^2 - 1\right)\left[\left(\theta_2^H - \overline{y}\right) - \tau\left(\theta_2^F - \overline{y}\right)\right]\right\} s_1^H L^H}{f_E + f_D + b s_1^H L^H}$$

which is decreasing in θ_2^F, and again independent of s_2^F. Due to Assumption 4.2 the following holds: $N_{X,2}^H > 0$, $N_{X,1}^H > 0$, and $N_1^H > 0$. Note that substituting the solutions into Foreign's balance of payments shows that the trade deficit in the amount of $-\left(\theta_1^F - \theta_2^F\right)s_1^F L^F < 0$ is equalized by a surplus in net factor payments $\left(\theta_1^F - \theta_2^F\right)s_1^F L^F > 0$.

4.A.2 Extension: Technological Differences

This section provides the formal appendix to the extension of differences in technologies across regions. We keep all our assumptions except that producers in Foreign cannot engage in FDI. We look at an equilibrium where all producers who sell to the middle and rich classes in both regions locate in both regions whereas all producers who exclude at least the middle class in Foreign locate in Home. Even though there is a homogeneous goods sector which demands labor, for full employment in Foreign there must be some producers of differentiated products located in Foreign if Assumption 4.2 holds.

Zero-profit Conditions

Only the profits of a producer who sells to the middle and rich classes in both regions and decides to locate in Foreign change. This producer has now only the option to export. In that case, she makes profits equal to

$$\pi_X^F = \left(z_2^F - bw^F\right)\left(s_1^F + s_2^F\right)L^F + \left(\tau z_2^F - \tau bw^F\right)\left(s_1^H + s_2^H\right)L^H - w^F\left(f_E + f_D\right).$$

In an equilibrium where this type of producer is active in both regions we must have $\pi_M^H = \pi_X^F$, which implies that relative wage rates w^H/w^F are determined by

$$\omega \equiv \frac{w^H}{w^F} = \frac{f_E + \tau b\left(s_1^H + s_2^H\right)L^H}{f_E + f_D + b\left(s_1^H + s_2^H\right)L^H}.$$

If relative wages fall short of the equilibrium wage rate ω we have $\pi_M^H > \pi_{X,2}^F$, and vice versa. Note that ω is larger than one if and only if $(\tau - 1)b > f_D/\left(s_1^H + s_2^H\right)L^H$, and is independent of s_2^F. Producers who choose to supply the middle and rich classes in Home but only the rich class in Foreign locate in Home if and only if $\phi_1 > \omega$, and vice versa. Similarly, all producers who decide to serve the rich class in Home and Foreign locate in Home if and only if $\phi_2 > \omega$, and vice versa. Finally, all producers who sell only to the rich class in Home locate in Home if and only if $\phi_3 > \omega$, and vice versa. Note that the ϕ's are defined as follows

$$\phi_1 \equiv \frac{f_E + f_D + \tau b\left(s_1^H + s_2^H\right)L^H + bs_1^F L^F}{f_E + f_D + b\left(s_1^H + s_2^H\right)L^H + \tau bs_1^F L^F}, \quad \phi_2 \equiv \frac{f_E + f_D + \tau bs_1^H L^H + bs_1^F L^F}{f_E + f_D + bs_1^H L^H + \tau bs_1^F L^F},$$
$$\phi_3 \equiv \frac{f_E + f_D + \tau bs_1^H L^H}{f_E + f_D + bs_1^H L^H}$$

where ϕ_1 is greater than one if and only if $\left(s_1^H + s_2^H\right)L^H > s_1^F L^F$, ϕ^2 is larger than one if and only if $s_1^H L^H > s_1^F L^F$, and ϕ_3 exceeds one if and only if $\tau > 1$. Note that Assumption 4.3 ensures that $\phi_1 > 1$ and $\phi_2 > 1$. One can show that $\phi_3 > \phi_2$ since $f_E + f_D + (\tau - 1)bs_1^H L^H > 0$, and $\phi_1 > \phi_2$ since $f_E + f_D + (\tau - 1)bs_1^F L^F > 0$. Hence, we assume that $\phi_2 > \omega$ holds, from which then follows that $\phi_1 > \omega$ and $\phi_3 > \omega$.

No producer located in Home has an incentive to deviate from its chosen mode of serving the foreign market if and only if

$$\frac{f_D}{s_1^F L^F} > (\tau\omega - 1)b > \frac{f_D}{\left(s_1^F + s_2^F\right)L^F}.$$

Due to our assumption producers located in Foreign have no choice but to export. Note that if producers located in Foreign could in principle engage in FDI, they won't do so if and only if

$$\frac{f_D}{s_1^H L^H} > \frac{f_D}{\left(s_1^H + s_2^H\right)L^H} > \left(\tau\frac{1}{\omega} - 1\right)b$$

which follows from comparing profits of Foreign producers from engaging in FDI with profits from exporting.

Resource and Budget Constraints

We have to modify the resource constraints as follows. The resource constraint in Home is now given by

$$\begin{aligned}\left(\theta_1^H s_1^H + \theta_2^H s_2^H + \theta_3^H s_3^H\right) L^H &= \left[\overline{x}\left(s_1^H + s_2^H\right) L^H + x_3^H s_3^H L^H\right] a \\ &+ M^H \left[f_E + f_D + b\left(s_1^H + s_2^H\right) L^H\right] \\ &+ N_{X,2}^H \left[f_E + f_D + b\left(s_1^H + s_2^H\right) L^H + \tau b s_1^F L^F\right] \\ &+ N_{X,1}^H \left[f_E + f_D + b s_1^H L^H + \tau b s_1^F L^F\right] \\ &+ N_1^H \left[f_E + f_D + b s_1^H L^H\right].\end{aligned}$$

Similarly, the resource constraint in Foreign is now determined by

$$\begin{aligned}\left(\theta_1^F s_1^F + \theta_2^F s_2^F + \theta_3^F s_3^F\right) L^F &= \left[\overline{x}\left(s_1^F + s_2^F\right) L^F + x_3^F s_3^F L^F\right] a \\ &+ N_X^F \left[f_E + f_D + b\left(s_1^F + s_2^F\right) L^F + \tau b\left(s_1^H + s_2^H\right) L^H\right] \\ &+ M^H \left[f_D + b\left(s_1^F + s_2^F\right) L^F\right].\end{aligned}$$

Note that the budget constraints do not change, except that $w^H = w^F$ no longer holds.

Prices

The price of food now differs across regions, and is given by

$$p_x^i = w^i a.$$

The willingness to pay of consumers is determined by

$$\begin{aligned}z_1^H &= \frac{w^H\left[f_E + f_D + b s_1^H L^H\right]}{s_1^H L^H} \\ z_2^H &= \frac{w^H\left[f_E + f_D + b\left(s_1^H + s_2^H\right) L^H + \tau b s_1^F L^F\right]}{\left(s_1^H + s_2^H\right) L^H + \tau s_1^F L^F} \\ z_1^F &= \frac{w^H\left[f_E + f_D + b s_1^H L^H + \tau b s_1^F L^F\right]}{\tau s_1^H L^H + s_1^F L^F} \\ z_2^F &= \frac{w^F\left[f_E + f_D + \tau b\left(s_1^H + s_2^H\right) L^H + b\left(s_1^F + s_2^F\right) L^F\right]}{\tau\left(s_1^H + s_2^H\right) L^H + \left(s_1^F + s_2^F\right) L^F}.\end{aligned}$$

Let us choose labor in Foreign as the numeraire, and set $w^F = 1$. Again, Assumption 4.2 implies that parameters are restricted so that $z_1^H \geq \tau z_1^F \geq \tau^2 z_2^H \geq \tau^3 z_2^F$ holds. Notice that $z_2^F > b$ holds since $(f_E + f_D) > 0$, and that $\partial z_2^F / \partial s_2^F < 0$.

Number of Producers

The number of producers that supply the middle class and the rich in Home and Foreign is determined by

$$M^H = \frac{\left(\theta_2^F - \overline{y}\right)\left[f_E + f_D + \tau b\left(s_1^H + s_2^H\right)L^H + b\left(s_1^F + s_2^F\right)L^F\right]}{z_2^F\left[f_E + \tau b\left(s_1^H + s_2^H\right)L^H\right]} - \frac{\sum_{k=1}^{2}\left(\theta_k^F - \overline{y}\right)s_k^F L^F}{f_E + \tau b\left(s_1^H + s_2^H\right)L^H}$$

which is increasing in θ_2^F if and only if $\left[f_E + f_D + b\left(s_1^F + s_2^F\right)L^F + \tau b\left(s_1^H + s_2^H\right)L^H\right]/z_2^F s_2^F L^F$ exceeds 1, and increasing in s_2^F if the following sufficient (but not necessary) condition holds

$$\varepsilon_{zs}\left[\frac{f_E + f_D + \tau b\left(s_1^H + s_2^H\right)L^H + b\left(s_1^F + s_2^F\right)L^F}{z_2^F s_2^F L^F}\right] > \frac{\left(z_2^F - b\right)}{z_2^F}$$

where we again assume $\varepsilon_{zs} \equiv -\left(\partial z_2^F / z_2^F\right) / \left(\partial s_2^F / s_2^F\right)$ to be greater than one.

The number of producers located in the Foreign that sells to all consumers above the income threshold is given by

$$N_X^F = \frac{\sum_{k=1}^{2}\left(\theta_k^F - \overline{y}\right)s_k^F L^F}{f_E + \tau b\left(s_1^H + s_2^H\right)L^H} - \frac{\left(\theta_2^F - \overline{y}\right)\left[f_D + b\left(s_1^F + s_2^F\right)L^F\right]}{z_2^F\left[f_E + \tau b\left(s_1^H + s_2^H\right)L^H\right]}$$

which is decreasing in θ_2^F if and only if $1 > \left[f_D + b\left(s_1^F + s_2^F\right)L^F\right]/z_2^F s_2^F L^F$, and decreasing in s_2^F if the following sufficient (but not necessary) condition holds

$$\varepsilon_{zs}\left[\frac{f_D + b\left(s_1^F + s_2^F\right)L^F}{z_2^F s_2^F L^F}\right] > \frac{\left(z_2^F - b\right)}{z_2^F}$$

where $\varepsilon_{zs} > 1$.

The total number of producers that serves the middle class and the rich in both regions can be written as

$$M^H + N_X^F = \frac{\left(\theta_2^F - \overline{y}\right)}{z_2^F} = \frac{\left(\theta_2^F - \overline{y}\right)\left[\tau\left(s_1^H + s_2^H\right)L^H + \left(s_1^F + s_2^F\right)L^F\right]}{f_E + f_D + \tau b\left(s_1^H + s_2^H\right)L^H + b\left(s_1^F + s_2^F\right)L^F}$$

which is increasing in θ_2^F and s_2^F. In an equilibrium, where $M^H > 0$ and $N_X^F > 0$ the following condition must hold

$$\begin{aligned} & \frac{f_E + f_D + b\left(s_1^F + s_2^F\right)L^F + \tau b\left(s_1^H + s_2^H\right)L^H}{z_2^F s_2^F L^F} > \frac{\left(\theta_1^F - \overline{y}\right)s_1^F L^F + \left(\theta_2^F - \overline{y}\right)s_2^F L^F}{\left(\theta_2^F - \overline{y}\right)s_2^F L^F} \\ > \; & \frac{f_D + b\left(s_1^F + s_2^F\right)L^F}{z_2^F s_2^F L^F} > 1 \end{aligned}$$

where the last inequality holds since repatriated profits from MNCs with HQ in Home are positive.

The number of producers who serve the middle and rich classes in Home and the rich class in Foreign is determined by

$$N_{X,2}^{H} = \frac{\left[\left(\theta_2^H - \overline{y}\right) w^H - \tau\left(\theta_2^F - \overline{y}\right)\right]\left[\left(s_1^H + s_2^H\right) L^H + \tau s_1^F L^F\right]}{w^H\left[f_E + f_D + b\left(s_1^H + s_2^H\right) L^H + \tau b s_1^F L^F\right]}$$

where w^H is determined above and $\overline{y} = a\overline{x}$. We see that $N_{X,2}^{H}$ is decreasing in θ_2^F, and independent of s_2^F since w^H/w^F is independent of s_2^F.

Similarly, the number of producers that supplies only the rich class in both regions is given by

$$N_{X,1}^{H} = \frac{\left\{\left[\left(\theta_1^F - \overline{y}\right) - \tau\left(\theta_2^H - \overline{y}\right) w^H\right] + \left(\tau^2 - 1\right)\left(\theta_2^F - \overline{y}\right)\right\}\left(\tau s_1^H L^H + s_1^F L^F\right)}{w^H\left[f_E + f_D + b s_1^H L^H + \tau b s_1^F L^F\right]}.$$

We see that $N_{X,1}^{H}$ rises in θ_2^F, and does not depend on s_2^F.

Eventually, the number of producers that only sells to the rich class in Home is given by

$$N_1^H = \frac{\left\{\left[\left(\theta_1^H - \overline{y}\right) w^H - \tau\left(\theta_1^F - \overline{y}\right)\right] + \left(\tau^2 - 1\right)\left[\left(\theta_2^H - \overline{y}\right) w^H - \tau\left(\theta_2^F - \overline{y}\right)\right]\right\} s_1^H L^H}{w^H\left[f_E + f_D + b s_1^H L^H\right]}$$

which is decreasing in θ_2^F, and also independent of s_2^F. Assumption 4.2 guarantees that all measures of producers are positive. Last, the balance of payments for Foreign is now given by

$$\begin{aligned} &- \left(N_{X,2}^H \tau z_2^H + N_{X,1}^H z_1^F\right) s_1^F L^F + N_X^H \tau z_2^F \left(s_1^H + s_2^H\right) L^H \\ &- \left[\left(z_2^F - b\right)\left(s_1^F + s_2^F\right) L^F - f_D\right] M^H \\ &= 0 \end{aligned}$$

where the first line denotes the trade balance, which now includes exports of the Foreign to the Home, and the second line net factor payments.

4.A.3 Extension: Homothetic Preferences

This section follows partly Brainard (1997) and Antras and Nunn (2009) in describing the baseline model with homothetic preferences. Consumer preferences are given by

$$U = x^{\beta} C^{1-\beta}$$

where $C = \left(\int_{\mathcal{J}} c(j)^{\frac{\sigma-1}{\sigma}} dj\right)^{\frac{\sigma}{\sigma-1}}$, where $\sigma > 1$, denotes the subutility which aggregates differentiated products into a composite good C. Cobb-Douglas preferences between food x and the composite good C imply that consumers spend a constant share β of their income y on food, and the rest on differentiated products (note that these preferences allow for two-stage budgeting).

Demand for differentiated product j can then be derived as follows. Consumers maximize subutility C subject to their budget constraint. The first-order conditions are determined by

$$\begin{aligned}\left(\int_{\mathcal{J}} c(j)^{\frac{\sigma-1}{\sigma}} dj\right)^{\frac{\sigma}{\sigma-1}-1} c(j)^{\frac{\sigma-1}{\sigma}-1} - \lambda p(j) &= 0 \\ (1-\beta)y - \int_{\mathcal{J}} p(j)c(j)dj &= 0.\end{aligned}$$

It is straightforward to show that if $c(j) \to 0$, marginal utility $\partial U/\partial c(j) \to \infty$. Hence, consumers will always purchase all differentiated products available on the market. We stick to our assumptions about the distribution of endowments except that we will now assume aggregate incomes to be the same across regions, i.e. $Y^i = Y$ for all $i, l = \{H, F\}$. The Marshallian demand function of a consumer belonging to group k who resides in region i for product j is given by

$$c_{ik}(j) = A_k^i p(j)^{-\sigma}$$

where $A_k^i \equiv (1-\beta)y_k^i / \left(\int_{\mathcal{J}} p(j)^{-(\sigma-1)} dj\right)$ is residual demand (which producers take as given). Note that $y_k^i = \theta_k^i w_k^i + v_k^i$, with $v_k^i = 0$ for all k in equilibrium. From the Marshallian demand curve it becomes evident that the distribution of income within region i does not matter for aggregate demand of good j, only aggregate income Y^i matters. To see this, sum individual demand $c_{ik}(j)$ over all groups k in region i, which yields

$$C_i(j) \equiv \sum_k s_k^i L^i c_{ik} = p(j)^{-\sigma} A^i = \frac{p(j)^{-\sigma}(1-\beta)Y^i}{\int_{\mathcal{J}} p(j)^{-(\sigma-1)} dj}$$

where $Y^i \equiv \sum_k s_k^i L^i y_k^i$ denotes aggregate income in region i, and $A^i \equiv \sum_k s_k^i L^i A_k^i$ aggregate residual demand. This is a property of homothetic preferences. The price elasticity of aggregate demand is given by $\sigma > 1$. The better substitutes products j and j' are for each other, i.e. $\sigma \to \infty$, the more strongly relative demand reacts to changes in relative prices. We also keep the assumptions about technology from the baseline model, and now turn to the profit maximization problem of a producer located in region i. She sells her product to all groups in both regions. If she decides to export, she makes profits equal to

$$\pi_X^i(j) = \left[p_{ii}(j) - bw^i\right] C_{ii}(j) + \left[p_{il}(j) - \tau bw^i\right] C_{il}(j) - w^i \left(f_E + f_D\right).$$

The first-order conditions imply that all differentiated products $j \in \mathcal{J}$ made region i are sold at prices $p_{ii} = \left(\frac{\sigma}{\sigma-1}\right) w^i b$ on the domestic market, and at prices $p_{il} = \left(\frac{\sigma}{\sigma-1}\right) \tau b w^i$ on the foreign market. Optimal profits can be written as follows

$$\pi_X^i = B^i + \tau^{-(\sigma-1)} B^l - w^i \left(f_E + f_D\right)$$

where $B^i \equiv (\sigma-1)^{\sigma-1}\sigma^{-\sigma}\left(bw^i\right)^{-(\sigma-1)} A^i$ for $i, l = \{H, F\}$. If producer j chooses to set up a foreign production facility instead, she earns profits given by

$$\pi_M^i(j) = \left[p_{ii}(j) - bw^i\right] C_{ii}(j) + \left[p_{il}(j) - bw^l\right] C_{il}(j) - \left[w^i\left(f_E + f_D\right) + w^l f_D\right].$$

The producer's optimality conditions imply that all products j manufactured in region i are sold at prices $p_{ii} = \left(\frac{\sigma}{\sigma-1}\right) w^i b$ in the domestic market. Optimal profits are then determined by

$$\pi_M^i = B^i + B^l - \left[w^i\left(f_E + f_D\right) + w^l f_D\right].$$

In a (symmetric) equilibrium with free entry all producers must earn zero profits, i.e. $\pi_M^H = \pi_M^F = 0$. This implies that relative wages equalize so that $w^H = w^F = 1$ if we choose labor in Home as the numeraire. In both cases, if we impose symmetry, i.e. $B^H = B^F = B$. In a symmetric equilibrium where all producers export, aggregate demand is equal to $B = \left(f_E + f_D\right) / \left(1 + \tau^{-(\sigma-1)}\right)$, whereas in a symmetric equilibrium where all producers engage in FDI, the aggregate demand level is determined by $\left(f_E + f_D\right)/2$. From the budget constraint of consumers, the Marshallian demand functions, and the aggregate demand level B, follows the number of producers in region i. In an equilibrium with pervasive exporting in each region, the number of exporters is determined by

$$N^i = \frac{(1/\sigma)(1-\beta)Y}{f_E + f_D}.$$

In an equilibrium with pervasive FDI, the number of multinationals is given by

$$M^i = \frac{(1/\sigma)(1-\beta)Y}{f_E + 2f_D}.$$

It becomes apparent here that the distribution within regions has no effect on the number of producers that export or are engaged in FDI. This is intuitive since with homothetic preferences the distribution within regions has no effect on aggregate demand. The equilibrium with pervasive exporting is an equilibrium if and only if no producer has an incentive to deviate and serve the foreign market by engaging in FDI. Thus, profits from setting up a foreign production plant, $2B - (f_E + 2f_D)$, must be negative. This implies that if and only if the following condition holds, no producer has an incentive to deviate

$$\frac{f_D}{f_E + f_D} > \frac{1 - \tau^{-(\sigma-1)}}{1 + \tau^{-(\sigma-1)}}.$$

This condition is the equivalent to condition (4.8) in the baseline model. We observe that market size, i.e. aggregate demand (number of people times per capita consumption), is absent in the condition above. The intuition is that higher aggregate demand induces more entry to the point where the demand level of an individual producer becomes independent of market size (see Antras and Nunn, 2009). In other words, a larger market has two opposing effects. On the one hand, it increases market size, which means higher profits for producers. On the other hand, a larger market implies fiercer competition and therefore lower profits. With CES preferences these two effects exactly offset each other in equilibrium (constant markups). Similarly, the equilibrium with pervasive FDI is an equilibrium if and only if no producer has an incentive to supply the foreign market through exports rather than engaging in FDI. It is straightforward to show that this is the case if the inequality above is reversed. However, this implies that a mixed equilibrium where some producers export and others engage in FDI only emerges if the condition above holds with equality. In sum, a mixed equilibrium occurs only in a knife-edge case where the number of exporters and multinationals is indeterminate (i.e. the parameter space for which this equilibrium exists has measure zero).

4.A.4 Extension: Skill versus Income Distribution

This section presents the formal model discussing the effects of skill versus income distribution. We will focus on a mixed equilibrium where both multinationals and exporters are active. It follows in part the factor endowment models of Markusen and Venables (2000), and Egger and Pfaffermayr (2005). We assume that the distribution of efficiency units of labor θ_k^i maps into a distribution of skills. In particular, we now assume θ_k^i to reflect the skills of group $k = \{1,2,3\}$ and region $i, l = \{H, F\}$. Suppose that in both regions group 1 supplies high-skilled, group 2 medium-skilled, and group 3 low-skilled labor.

We keep the assumption about homothetic consumer preferences. Thus, aggregate demand for food in region i is given by $X^i = \beta Y^i / p_x^i$, where $Y^i = \sum_k s_k^i L^i \theta_k^i w_k^i$ in equilibrium. Aggregate demand for any differentiated product j produced and sold in region i is given by $C_{ii} = (1-\beta) Y^i p_{ii}^{-\sigma} / P^i$, and by $C_{il}/\tau = (1-\beta) Y^l \left(\tau p_{ll}\right)^{-\sigma} / P^l$ if the product is sold in region l, with $P^i = \left(M_i + N_i + M_l\right) p_{ii}^{-(\sigma-1)} + N_l \left(\tau p_{ll}\right)^{-(\sigma-1)}$.

However, on the production side we make the assumption that producers combine different skills to produce using the same technology across regions. The differentiated product sector is more skill-intensive combining medium and high-skilled labor than the food sector, which uses only low-skilled labor. For simplicity, we assume that the food sector uses a units of low-skilled labor to produce 1 unit of output. Food is not traded and produced under conditions of perfect competition, which implies that its price is equal to marginal costs, $p_x^i = a w_3^i$. In the differentiated sector, producers use f_E units of high-skilled labor to create a new product, and use f_D, and b units of medium-skilled labor to set up a new production facility, and produce 1 unit of output, respectively. We assume that multinational producers employ high-skilled labor in their region of origin, and medium-skilled labor in the region of production. They operate under conditions of monopolistic competition. Optimal monopoly prices for all products made and sold in region i are given by $p_{ii} = \left(\frac{\sigma}{\sigma-1}\right) b w_2^i$, and $p_{il} = \left(\frac{\sigma}{\sigma-1}\right) \tau b w_2^i$ for exported products. Optimal profits of multinational, and exporters, respectively, in region i are determined by

$$
\begin{aligned}
\pi_M^i &= (\sigma-1)^{-1} \left(b w_2^i\right) C_{ii} + (\sigma-1)^{-1} \left(b w_2^l\right) C_{ll} - w_1^i f_E - w_2^i f_D - w_2^l f_D \\
\pi_X^i &= (\sigma-1)^{-1} \left(b w_2^i\right) C_{ii} + (\sigma-1)^{-1} \left(\tau b w_2^i\right) C_{il} - w_1^i f_E - w_2^i f_D
\end{aligned}
$$

which must both be equal to zero in a free entry equilibrium where multinationals and exporters co-exist. Consider for the moment a symmetric (mixed) equilibrium. Free entry and zero-profits in equilibrium, i.e. $\pi_X^i = \pi_M^i = 0$, imply

$$\frac{1-\tau^{-(\sigma-1)}}{1+\tau^{-(\sigma-1)}} = \frac{w_2 f_D}{w_1 f_E + w_2 f_D}$$

where the right-hand side is decreasing in w_1/w_2. Only multinationals are active in equilibrium if and only if the left-hand side above exceeds the right-hand side, and vica versa. Thus, if the relative wage rate w_1/w_2 is low, exporting is more attractive than engaging in FDI, ceteris paribus. Notice that if the same type of labor is used to create new products and set up new factories, i.e. if $w_1/w_2 = 1$, this condition reduces to the knife-edge condition of the previous section. However, in this model the parameter space for which a mixed equilibrium exists has non-zero measure. Finally, labor market clearing in region i demands that

$$\begin{aligned} s_1^i L^i \theta_1^i &= \left(M^i + N^i\right) f_E \\ s_2^i L^i \theta_2^i &= M^i \left(f_D + C_{ii} b\right) + N^i \left(f_D + C_{ii} b + C_{il} b\right) + M^l \left(f_D + C_{ii} b\right) \\ s_3^i L^i \theta_3^i &= X^i a \end{aligned}$$

holds. Note that the balance of payments is implied by the zero-profit conditions, labor market clearing and budget constraints (this is an accounting identity). Due to Walras' law we are free to drop the market clearing condition for food in Foreign, and set its price equal to one, i.e. $p_x^F = 1$. Hence, we have 23 (non-linear) equations in 23 unknowns (C_{ii}, C_{il}), Y^i, (p_{ii}, p_{ll}), p_x^i, P^i, $\left(w_1^i, w_2^i, w_3^i\right)$, $\left(M^i, N^i\right)$, X^i for $i, l = \{H, F\}$, and where $p_x^F = 1$. It is not possible to solve this model analytically. Thus, we simulate the model to perform comparative statics by looking at small changes in parameters.

The simulations in Figure A.1 below illustrate how a decrease in the average labor endowment θ_2^F of medium-skilled in Foreign affects the number of multinationals relative to exporters in Home if we increase (i) θ_1^F, and (ii) θ_3^F so that the total skill endowment $\sum_k s_k^F L^F \theta_k^F$ is constant. The simulations in Figure A.2 show how the number of MNCs with HQ in Home changes when we decrease the population share s_2^F of the middle-skilled in Foreign, and increase (iii) s_1^F and (iv) s_3^F such that aggregate skill endowment in Foreign is constant. In our simulations we choose parameters such that high/middle/low skills translate into high/middle/low per capita incomes. Figure A.1 shows that in case (i) there is a positive link between per capita income of the middle class in Foreign and the number of MNCs with HQ in Home, whereas in case (ii) the link is negative. In Figure A.2 we see that a shrinking size of the middle class leads in both case (iii) and (iv) to a decline in the number of MNCs with HQ in Home. The basic intuition is that if the supply of medium-skilled labor is relatively low, the wage rate for medium-skilled labor is high (i.e. w_1/w_2 is low), ceteris paribus. This means that exporting becomes more attractive relative to engaging in FDI, ceteris paribus.

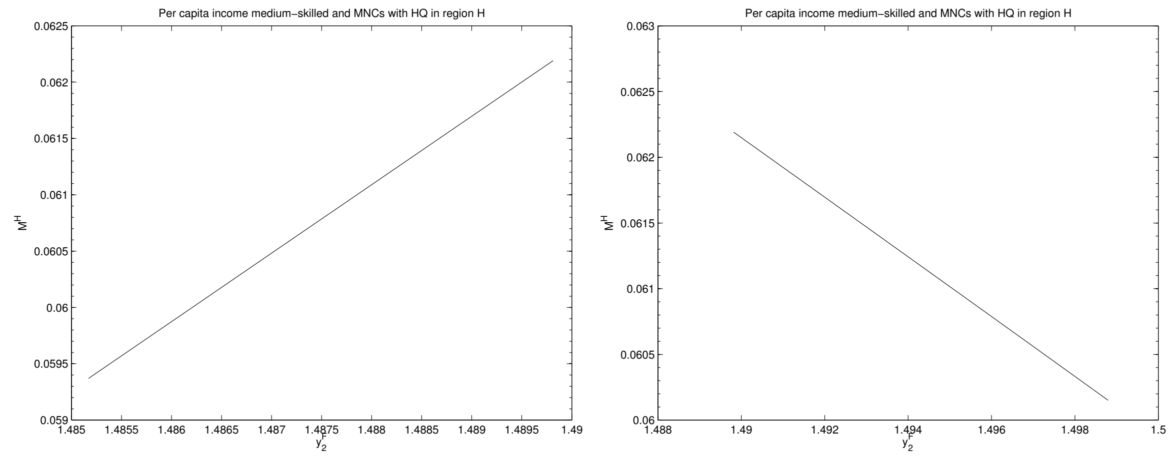


Figure A.1: Effect of changes in Foreign's skill distribution on MNCs with HQ in Home: case (i) on left-hand side, case (ii) on right-hand side. Parameter values: $\beta = 0.1$, $\sigma = 2$, $\tau = 1.13$, $a = b = 2$, $f_E = 2.5$, $f_D = 0.33$, $s_1^i = 0.2$, $s_2^i = 0.5$, $s_3^i = 0.3$, $L^i = 1$, and $\theta_k^i = 1$ as starting values, for all i, k.

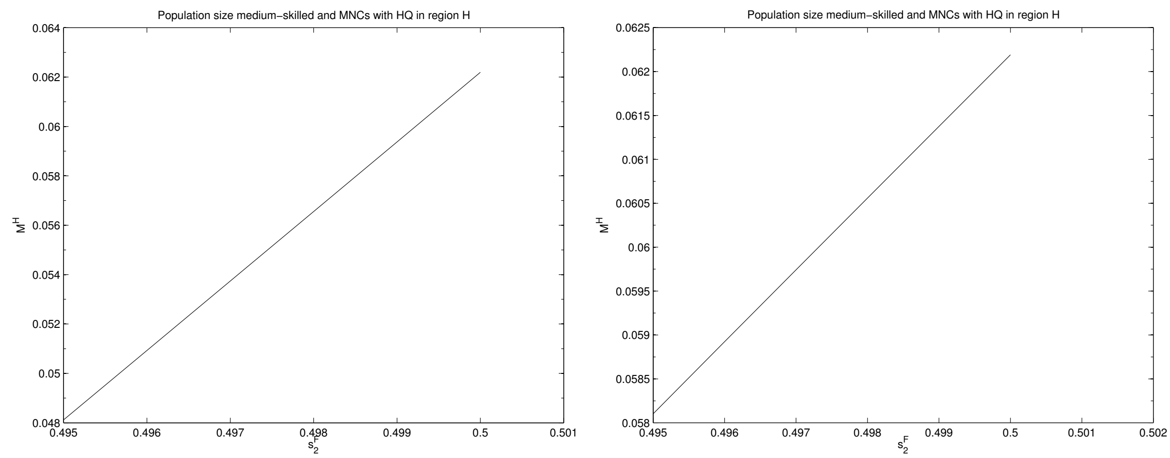


Figure A.2: Effect of changes in Foreign's skill distribution on MNCs with HQ in Home: case (iii) on left-hand side, case (iv) on right-hand side. Parameter values: $\beta = 0.1$, $\sigma = 2$, $\tau = 1.13$, $a = b = 2$, $f_E = 2.5$, $f_D = 0.33$, $s_1^i = 0.2$, $s_2^i = 0.5$, and $s_3^i = 0.3$ as starting values, $L^i = 1$, and $\theta_k^i = 1$ for all i, k.

4.A.5 Empirics

This appendix provides detailed information on the income data used, summary statistics, and some robustness checks.

Income Data

Bernasconi (2013) constructs discrete empirical income distributions for 94 countries for the time period 1997-2007 using income shares of deciles and quintiles from UNU-WIDER (2008) and GDP per capita from Heston et al. (2012). See Bernasconi (2013) for a detailed description of how she transforms the inequality data into income distributions. In short, she assigns an average income level to each decile (quintile), and then redistributes the corresponding area uniformly on an income interval. However, she does not make parametric assumptions about the functional form of the income distribution across income intervals. Finally, the resulting densities are divided into common income intervals on USD $1, 5000, \ldots, 145000, 150000$. Note that Bernasconi (2013) assumes that the income shares do not change over time (i.e. the Lorenz curve is constant over time). However, the income distributions change over time as average incomes of deciles (quintiles) change. She selects data on the basis of consistency (e.g. income versus expenditure inequality) and quality in order to increase data consistency (see Bernasconi, 2013). This assumption implies that there is less variation over time but makes the data useful to study cross-section data.

Note that we adjust the empirical income distributions such that the aggregate income computed from the distributions is close to GDP (Y) from Heston et al. (2012). We calculate the mean income of bin l as the generalized mean $\widetilde{y}_l = \left(\underline{y}_l^{\theta}/2 + \overline{y}_l^{\theta}/2\right)^{1/\theta}$ where $\underline{y}_l$ denotes the lower bound of bin i and $\overline{y}_l$ the upper bound, and θ is a non-zero real number. For each country (and year) we choose θ such that aggregate income, i.e. $\widetilde{Y} = \sum_l \widetilde{y}_l P_l$ where P_l denotes the number of people in bin l, is at most 35 percent above or below Y. We compute $\widetilde{Y}$ by iterating $\theta = 1, 1/2, \ldots, 1/1000$ until our convergence criteria, i.e. $0.65 < \widetilde{Y}/Y < 1.35$, is satisfied. In less than 2 percent of all cases, a thousand steps are not enough to converge. In these cases the last iteration step is taken. In general, the chosen θ is such that the computed mean incomes deviate more from the simple arithmetic averages in low income bins than in high income bins. This suggests that the discrepancy between $\widetilde{Y}$ and Y is because $\widetilde{Y}$ overestimates the mean incomes in the lower income bins. Note that if $\theta = 1$ the generalized mean collapses to the simple arithmetic mean. However, since we control for GDP from Heston et al. (2012) in the regressions this issue is of second order.

Summary Statistics and Robustness

The list of countries can be found in Table A.1. Summary statistics about the dependent and independent variables are given in Table A.2. As an example, Table A.3 shows the mean over 1997-2007 of the log of the market size measure for the income thresholds $\{\underline{y}, \overline{y}\} = \{5000, 20000\}$ for all host countries in the sample. Table A.4 and A.5 show the estimation of equation (4.9) separately for each year 1997-2007, with a lower threshold of USD 5,000. The results for a lower threshold of USD 10,000 are similar.

Table A.1: Country list

Parent countries	**Host countries**		
Australia	Argentina	Honduras	Turkey
Austria	Australia	Hungary	Uganda
Benelux	Austria	Indonesia	United Kingdom
Canada	Bangladesh	Iran, Islamic Rep.	United States
Chile	Benelux	Ireland	Venezuela, Bolivar Rep.
Denmark	Benin	Israel	Vietnam
Finland	Bolivia	Italy	Yemen, Rep.
Fmr Czechoslovakia	Brazil	Jamaica	Zambia
France	Bulgaria	Korea, Rep.	
Germany	Cambodia	Mexico	
Greece	Canada	Morocco	
Hungary	Chile	Nepal	
Iceland	Colombia	Netherlands	
Ireland	Costa Rica	Nicaragua	
Israel	Denmark	Norway	
Italy	Dominican Rep.	Pakistan	
Japan	Ecuador	Panama	
Korea, Rep.	Egypt, Arab Rep.	Paraguay	
Netherlands	El Salvador	Peru	
New Zealand	Finland	Philippines	
Norway	Fmr Czechoslovakia	Poland	
Poland	Fmr USSR	Portugal	
Portugal	Fmr Yugoslavia	Romania	
Spain	France	South Africa	
Sweden	Germany	Spain	
Switzerland	Greece	Sri Lanka	
Turkey	Guatemala	Sweden	
United Kingdom	Guyana	Tanzania	
United States	Haiti	Thailand	

Notes: Australia includes external territories, Benelux includes Belgium and Luxembourg, Denmark includes Faroe Islands, France includes Andorra and Monaco, Israel includes West Bank, Italy includes San Marino and Vatical City, New Zealand includes Norfolk, Cocos and Christmas Islands, Switzerland includes Liechtenstein, Indonesia includes Macao, South Africa includes Rep. of South Africa, Botswana, Lesotho, Namibia, and Swaziland.
We drop Mexico as a parent country from the sample since they report only 4 outward FDI positions over 1997-2007.

Table A.2: Summary statistics, 1997-2007

	min	p25	p50	mean	p75	max	sd	N
FDI position$_{ijt}$ (m USD, PPP)	0	0	195	5910	2506	4.00e+05	22868	11817
$\log(Y\,[y\|y \leq 5,000]_{jt})$	5.637	7.698	8.909	8.924	10.120	12.831	1.667	11817
$\log(P\,[y\|y \leq 5,000]_{jt})$	4.726	7.010	8.399	8.443	9.981	11.914	1.885	11817
$\log(Y\,[y\|y \leq 10,000]_{jt})$	6.665	9.311	10.218	10.263	11.514	14.003	1.592	11817
$\log(P\,[y\|y \leq 10,000]_{jt})$	5.415	7.846	8.926	9.013	10.253	12.405	1.721	11817
$\log(Y\,[y\|10,000 \leq y > 5,000]_{jt})$	0.000	8.680	9.869	9.728	11.131	13.800	1.986	11817
$\log(P\,[y\|10,000 \leq y > 5,000]_{jt})$	0.000	6.688	7.854	7.755	9.141	11.786	1.864	11817
$\log(Y\,[y\|15,000 \leq y > 5,000]_{jt})$	0.000	9.249	10.998	10.558	11.705	14.326	1.952	11817
$\log(P\,[y\|15,000 \leq y > 5,000]_{jt})$	0.000	7.032	8.770	8.344	9.437	12.091	1.813	11817
$\log(Y\,[y\|20,000 \leq y > 5,000]_{jt})$	0.000	10.087	11.434	11.093	12.414	14.505	1.921	11817
$\log(P\,[y\|20,000 \leq y > 5,000]_{jt})$	0.000	7.616	9.015	8.703	9.897	12.185	1.757	11817
$\log(Y\,[y\|25,000 \leq y > 5,000]_{jt})$	0.000	10.570	11.620	11.430	12.646	14.652	1.945	11817
$\log(P\,[y\|25,000 \leq y > 5,000]_{jt})$	0.000	7.874	9.191	8.919	10.156	12.255	1.746	11817
$\log(Y\,[y\|30,000 \leq y > 5,000]_{jt})$	0.000	10.837	11.846	11.652	12.771	14.721	1.984	11817
$\log(P\,[y\|30,000 \leq y > 5,000]_{jt})$	0.000	8.118	9.270	9.051	10.259	12.269	1.752	11817
$\log(Y\,[y\|35,000 \leq y > 5,000]_{jt})$	0.000	10.942	12.002	11.799	12.956	14.996	2.013	11817
$\log(P\,[y\|35,000 \leq y > 5,000]_{jt})$	0.000	8.217	9.330	9.133	10.298	12.284	1.758	11817
$\log(Y\,[y\|15,000 \leq y > 10,000]_{jt})$	0.000	8.743	10.099	9.726	11.229	13.643	2.249	11817
$\log(P\,[y\|15,000 \leq y > 10,000]_{jt})$	0.000	6.219	7.573	7.283	8.703	11.118	1.935	11817
$\log(Y\,[y\|20,000 \leq y > 10,000]_{jt})$	0.000	9.732	10.818	10.518	11.996	13.970	2.315	11817
$\log(P\,[y\|20,000 \leq y > 10,000]_{jt})$	0.000	7.025	8.118	7.916	9.341	11.361	1.962	11817
$\log(Y\,[y\|25,000 \leq y > 10,000]_{jt})$	0.000	10.325	11.322	10.952	12.243	14.271	2.387	11817
$\log(P\,[y\|25,000 \leq y > 10,000]_{jt})$	0.000	7.463	8.494	8.234	9.520	11.513	2.003	11817
$\log(Y\,[y\|30,000 \leq y > 10,000]_{jt})$	0.000	10.570	11.568	11.225	12.570	14.670	2.448	11817
$\log(P\,[y\|30,000 \leq y > 10,000]_{jt})$	0.000	7.730	8.685	8.419	9.624	11.671	2.038	11817
$\log(Y\,[y\|35,000 \leq y > 10,000]_{jt})$	0.000	10.786	11.759	11.402	12.697	14.958	2.487	11817
$\log(P\,[y\|35,000 \leq y > 10,000]_{jt})$	0.000	7.855	8.784	8.530	9.783	11.858	2.060	11817
$\log(Y[y\|y > 10,000]_{jt})$	0.000	11.553	12.490	12.077	13.560	16.392	2.668	11817
$\log(P[y\|y > 10,000]_{jt})$	0.000	8.247	9.041	8.831	10.106	12.535	2.124	11817
$\log(Y[y\|y > 15,000]_{jt})$	0.000	11.125	12.336	11.705	13.426	16.371	3.182	11817
$\log(P[y\|y > 15,000]_{jt})$	0.000	7.656	8.809	8.351	9.764	12.455	2.461	11817
$\log(Y[y\|y > 20,000]_{jt})$	0.000	10.845	12.194	11.247	13.288	16.334	3.791	11817
$\log(P[y\|y > 20,000]_{jt})$	0.000	7.207	8.419	7.869	9.614	12.345	2.829	11817
$\log(Y[y\|y > 25,000]_{jt})$	0.000	10.645	12.088	10.913	13.173	16.281	4.092	11817
$\log(P[y\|y > 25,000]_{jt})$	0.000	6.918	8.194	7.501	9.322	12.211	2.969	11817
$\log(Y[y\|y > 30,000]_{jt})$	0.000	10.253	11.909	10.547	12.997	16.213	4.320	11817
$\log(P[y\|y > 30,000]_{jt})$	0.000	6.582	7.912	7.108	9.007	12.061	3.064	11817
$\log(Y[y\|y > 35,000]_{jt})$	0.000	9.718	11.738	9.993	12.912	16.132	4.739	11817
$\log(P[y\|y > 35,000]_{jt})$	0.000	5.834	7.626	6.621	8.722	11.897	3.274	11817
log(GDP pc$_{jt}$)	6.507	8.554	9.162	9.193	10.295	10.855	1.109	11817
log(GDP$_{jt}$)	3.188	7.101	7.796	7.789	8.870	11.785	1.581	11817
log(remoteness$_{jt}$)	8.609	8.665	8.827	8.915	9.128	9.441	0.260	11817
log(urban pop$_{jt}$)	2.493	3.793	4.166	4.002	4.328	4.501	0.467	11817
log(openness$_{jt}$)	2.863	3.934	4.166	4.194	4.477	5.165	0.440	11817
log(tax rate$_{jt}$)	2.795	3.628	3.842	3.827	4.003	4.683	0.321	11817
log(distance$_{jt}$)	5.050	7.538	8.747	8.349	9.155	9.875	1.007	11817
log(squared skill diff$_{jt}$)	-12.296	0.946	2.718	2.085	3.769	5.153	2.280	11817
dummy common language$_{ijt}$	0.000	0.000	0.000	0.162	0.000	1.000	0.369	11817
dummy regional trade agreement$_{ijt}$	0.000	0.000	0.000	0.090	0.000	1.000	0.286	11817
dummy customs union$_{ijt}$	0.000	0.000	0.000	0.219	0.000	1.000	0.414	11817
dummy colonial ties$_{ijt}$	0.000	0.000	0.000	0.029	0.000	1.000	0.168	11817
log(percent no schooling$_{jt}$)	-2.310	0.911	1.851	1.748	2.743	4.392	1.461	11817
log(percent primary schooling$_{jt}$)	0.453	2.382	2.945	2.701	3.234	3.806	0.756	11817
log(percent secondary schooling$_{jt}$)	-0.066	2.643	3.075	2.948	3.481	3.927	0.731	11817
log(percent tertiary schooling$_{jt}$)	-1.390	1.494	2.198	1.949	2.575	3.404	0.983	11817

Table A.3: Mean of (log) market size measure by host country for $\{\underline{y}, \overline{y}\} = \{5000, 20000\}$

Country	$\log(Y_{low})$	$\log(P_{low})$	$\log(Y_{middle})$	$\log(P_{middle})$	$\log(Y_{high})$	$\log(P_{high})$
South Africa	7.923	10.386	11.639	9.475	13.123	8.918
Morocco	10.143	9.892	11.253	9.006	11.524	8.194
Egypt, Arab Rep.	11.155	10.904	11.735	9.476	12.515	8.792
Benin	8.492	8.791	8.785	6.810	0.000	0.000
Uganda	9.763	10.113	10.006	7.744	0.000	0.000
Tanzania	10.187	10.486	5.769	4.422	0.000	0.000
Zambia	8.147	9.167	9.263	6.869	5.519	3.631
Canada	8.126	7.210	11.587	8.942	13.819	10.022
United States	10.468	9.551	13.715	11.107	16.212	12.258
Argentina	10.072	9.156	12.324	9.974	12.848	8.878
Bolivia	9.102	8.851	9.337	7.072	9.943	6.473
Brazil	11.123	11.546	13.072	10.848	14.269	10.205
Chile	9.273	8.357	11.322	9.000	12.237	8.043
Colombia	10.405	10.153	11.565	9.314	12.433	8.458
Ecuador	9.383	9.132	10.220	7.940	10.999	7.148
Mexico	11.523	10.606	12.987	10.735	13.676	9.706
Paraguay	8.624	8.373	9.384	7.157	9.748	6.207
Peru	10.064	9.813	11.039	8.812	11.544	7.824
Venezuela, Bolivar Rep.	9.920	9.003	11.714	9.413	12.124	8.285
Costa Rica	8.154	7.238	9.842	7.557	10.322	6.456
El Salvador	8.620	8.369	9.622	7.399	9.965	6.350
Guatemala	9.245	8.994	10.185	7.910	10.855	7.036
Honduras	8.879	8.627	9.081	6.827	9.366	6.034
Nicaragua	8.132	8.431	8.746	6.352	8.063	4.948
Dominical Rep.	9.263	8.347	10.440	8.186	10.783	6.942
Haiti	8.604	8.904	9.161	6.659	5.208	2.917
Jamaica	7.530	7.279	8.988	6.762	9.831	5.879
Guyana	6.229	6.424	7.462	4.976	1.308	0.615
Panama	7.671	7.420	9.183	6.932	9.840	5.955
Israel	7.025	6.109	10.605	8.063	11.689	7.938
Iran, Islamic Rep.	8.098	10.033	12.524	10.286	13.095	9.354
Turkey	10.457	10.443	12.324	10.117	12.774	9.263
Yemen, Rep.	10.145	9.894	0.000	0.000	0.000	0.000
Bangladesh	11.955	11.703	12.261	10.067	0.000	0.000
Cambodia	9.267	9.333	9.630	7.255	8.356	5.427
Sri Lanka	9.128	9.658	10.196	7.848	11.533	7.561
Indonesia	12.135	11.883	13.199	10.912	13.214	9.956
Korea, Rep.	9.254	8.337	12.593	10.029	13.510	9.894
Nepal	9.272	9.986	10.892	8.454	0.000	0.000
Pakistan	11.985	11.734	12.795	10.385	1.531	1.075
Philippines	11.354	11.103	11.712	9.471	12.131	8.701
Thailand	10.706	10.455	12.166	9.893	12.915	9.027
Vietnam	11.504	11.126	11.672	9.358	10.902	7.758
Benelux	6.831	5.916	10.364	7.677	12.769	9.015
Denmark	6.544	5.630	10.010	7.368	11.982	8.163
France	8.775	7.858	12.578	9.895	14.261	10.579
Germany	8.978	8.062	12.612	9.933	14.648	10.975
Greece	7.509	6.593	11.032	8.442	12.227	8.536
Ireland	5.867	4.955	9.331	6.670	11.768	7.968
Italy	8.926	8.010	12.524	9.882	14.214	10.470
Netherlands	7.145	6.229	10.491	7.823	13.156	9.480
Portugal	7.614	6.698	11.229	8.699	11.970	8.197
Spain	8.576	7.660	12.201	9.557	13.786	10.055
United Kingdom	8.809	7.892	12.522	9.852	14.283	10.539
Austria	6.461	5.547	9.802	7.136	12.474	8.797
Finland	6.135	5.222	9.904	7.198	11.822	8.189
Norway	5.722	4.810	8.606	6.036	12.210	8.293
Sweden	6.681	5.767	10.430	7.724	12.402	8.738
Bulgaria	8.703	7.787	10.586	8.384	9.818	6.396
Fmr Czechoslovakia	7.761	6.846	11.866	9.297	11.736	8.195
Hungary	7.503	6.588	11.426	8.907	11.053	7.529
Poland	9.568	8.652	12.564	10.204	12.262	8.632
Romania	9.736	8.820	11.695	9.508	10.702	7.363
Fmr Yugoslavia	8.450	7.534	11.575	9.300	10.544	7.138
Fmr USSR	12.276	11.360	14.230	12.006	13.733	10.207
Australia	7.956	7.040	10.921	8.330	13.477	9.567

Notes: low is $y|y \leq \underline{y}$, middle is $y|\overline{y} \leq y > \underline{y}$, high is $y|y > \overline{y}$. Y is in millions, and P in thousands.

Table A.4: Outward FDI positions (29 OECD countries 1997-2007, PPML) with lower threshold $\underline{y} = 5000$

		$\log(FDI_{ijt})$	$\log(FDI_{ijt})$	$\log(FDI_{ijt})$	$\log(FDI_{ijt})$	$\log(FDI_{ijt})$	$\log(FDI_{ijt})$
		$\overline{y} = 10,000$	$\overline{y} = 15,000$	$\overline{y} = 20,000$	$\overline{y} = 25,000$	$\overline{y} = 30,000$	$\overline{y} = 35,000$
1997	$\log(Y_{middle,jt})$	1.658	0.806	2.146	2.987**	2.846***	2.128**
		(1.517)	(1.337)	(1.347)	(1.296)	(1.053)	(0.936)
	$\log(P_{middle,jt})$	-2.270	-1.375	-2.956*	-3.874**	-3.722***	-2.863**
		(1.832)	(1.658)	(1.681)	(1.583)	(1.303)	(1.173)
1998	$\log(Y_{middle,jt})$	1.644	1.396	2.626*	3.568***	3.059***	1.968*
		(1.632)	(1.415)	(1.373)	(1.321)	(1.119)	(1.008)
	$\log(P_{middle,jt})$	-2.229	-2.071	-3.515**	-4.543***	-3.944***	-2.652**
		(1.975)	(1.765)	(1.709)	(1.587)	(1.374)	(1.270)
1999	$\log(Y_{middle,jt})$	2.428	1.862	2.904**	3.586***	3.271***	2.290**
		(1.730)	(1.535)	(1.333)	(1.229)	(1.142)	(1.041)
	$\log(P_{middle,jt})$	-3.182	-2.675	-3.894**	-4.629***	-4.203***	-3.057**
		(2.080)	(1.883)	(1.648)	(1.503)	(1.373)	(1.281)
2000	$\log(Y_{middle,jt})$	2.914*	2.403*	3.173***	3.643***	3.615***	2.571**
		(1.623)	(1.457)	(1.216)	(1.144)	(1.206)	(1.003)
	$\log(P_{middle,jt})$	-3.674*	-3.254*	-4.155***	-4.637***	-4.470***	-3.298***
		(1.927)	(1.758)	(1.485)	(1.400)	(1.354)	(1.213)
2001	$\log(Y_{middle,jt})$	6.640***	5.535***	5.473***	6.480***	7.183***	4.864***
		(1.665)	(1.303)	(0.995)	(1.817)	(1.592)	(1.802)
	$\log(P_{middle,jt})$	-8.093***	-7.079***	-6.888***	-7.644***	-7.496***	-5.492***
		(2.022)	(1.606)	(1.232)	(1.784)	(1.357)	(1.448)
2002	$\log(Y_{middle,jt})$	3.495*	2.999*	3.640***	3.738***	3.569***	2.659***
		(1.855)	(1.696)	(1.246)	(1.024)	(1.056)	(0.966)
	$\log(P_{middle,jt})$	-4.320**	-3.925*	-4.639***	-4.664***	-4.341***	-3.340***
		(2.195)	(2.027)	(1.498)	(1.237)	(1.200)	(1.148)
2003	$\log(Y_{middle,jt})$	3.037**	2.078	3.122***	3.660***	3.589***	2.790***
		(1.320)	(1.336)	(1.116)	(1.003)	(1.159)	(0.974)
	$\log(P_{middle,jt})$	-3.627**	-2.711*	-3.884***	-4.384***	-4.194***	-3.351***
		(1.546)	(1.558)	(1.322)	(1.182)	(1.261)	(1.140)
2004	$\log(Y_{middle,jt})$	3.084*	2.401	3.466***	3.605***	3.523***	2.703***
		(1.601)	(1.510)	(1.290)	(1.059)	(1.083)	(1.013)
	$\log(P_{middle,jt})$	-3.702*	-3.128*	-4.374***	-4.452***	-4.208***	-3.323***
		(1.905)	(1.814)	(1.571)	(1.302)	(1.259)	(1.222)
2005	$\log(Y_{middle,jt})$	2.541**	1.659	2.854***	2.984***	3.750**	2.540**
		(1.182)	(1.204)	(1.096)	(0.934)	(1.858)	(1.061)
	$\log(P_{middle,jt})$	-2.975**	-2.167	-3.555***	-3.609***	-4.047**	-2.913**
		(1.405)	(1.446)	(1.345)	(1.157)	(1.654)	(1.158)
2006	$\log(Y_{middle,jt})$	1.912	1.365	2.585**	2.650**	2.907***	2.345**
		(1.213)	(1.288)	(1.226)	(1.085)	(1.063)	(0.987)
	$\log(P_{middle,jt})$	-2.216	-1.783	-3.246**	-3.279**	-3.419***	-2.794**
		(1.430)	(1.546)	(1.502)	(1.329)	(1.245)	(1.186)
2007	$\log(Y_{middle,jt})$	1.775	1.379	2.360*	2.619**	3.083***	2.518**
		(1.294)	(1.285)	(1.238)	(1.185)	(1.185)	(1.031)
	$\log(P_{middle,jt})$	-2.120	-1.836	-3.052**	-3.336**	-3.706***	-3.078**
		(1.521)	(1.535)	(1.510)	(1.440)	(1.363)	(1.224)

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered (by host country) standard errors in parentheses. Omitting rich class. Low is $y|y \leq \underline{y}$, middle is $y|\overline{y} \leq y > \underline{y}$, high is $y|y > \overline{y}$ for j, t. Controls: host remoteness, host urban concentration, host trade openness, host corporate tax, distance, squared skill difference, percent of population that completed no, primary, secondary, and tertiary schooling, dummies for common language, regional trade agreement, customs union, colonial relationship, parent country, and host region. Sample sizes (year): 661 (1997), 781 (1998), 832 (1999), 895 (2000), 992 (2001), 1013 (2002), 1175 (2003), 1278 (2004), 1264 (2005), 1424 (2006), 1493 (2007).

Table A.5: Outward FDI positions (29 OECD countries 1997-2007, PPML) with lower threshold $\underline{y} = 5000$

		$\log(FDI_{ijt})$	$\log(FDI_{ijt})$	$\log(FDI_{ijt})$	$\log(FDI_{ijt})$	$\log(FDI_{ijt})$	$\log(FDI_{ijt})$
		$\overline{y} = 10,000$	$\overline{y} = 15,000$	$\overline{y} = 20,000$	$\overline{y} = 25,000$	$\overline{y} = 30,000$	$\overline{y} = 35,000$
1997	$\log(Y_{middle,jt})$	3.046	-5.684**	1.638	3.568**	5.763***	2.821***
		(4.483)	(2.621)	(1.343)	(1.480)	(1.338)	(0.847)
	$\log(P_{middle,jt})$	-3.796	5.315*	-2.269	-4.579**	-7.469***	-3.801***
		(4.738)	(2.755)	(1.616)	(1.793)	(1.682)	(1.083)
1998	$\log(Y_{middle,jt})$	2.975	-3.885*	2.723**	5.686***	4.410***	2.998***
		(5.466)	(2.313)	(1.267)	(1.335)	(1.038)	(0.910)
	$\log(P_{middle,jt})$	-3.821	3.380	-3.520**	-7.168***	-5.884***	-4.073***
		(5.737)	(2.461)	(1.552)	(1.583)	(1.346)	(1.174)
1999	$\log(Y_{middle,jt})$	2.499	-1.871	3.388***	6.074***	4.827***	4.040***
		(5.204)	(2.244)	(1.221)	(1.209)	(0.973)	(1.058
	$\log(P_{middle,jt})$	-3.439	1.076	-4.510***	-7.760***	-6.423***	-5.416***
		(5.454)	(2.402)	(1.492)	(1.467)	(1.249)	(1.353)
2000	$\log(Y_{middle,jt})$	1.639	-0.749	3.719***	5.713***	4.639***	4.789***
		(5.655)	(2.160)	(1.171)	(1.031)	(0.878)	(1.126)
	$\log(P_{middle,jt})$	-2.709	-0.237	-4.906***	-7.270***	-6.002***	-6.291***
		(5.882)	(2.300)	(1.389)	(1.260)	(1.118)	(1.446)
2001	$\log(Y_{middle,jt})$	7.944*	3.133*	5.501***	7.439***	5.967***	5.546***
		(4.078)	(1.702)	(0.955)	(1.493)	(0.645)	(0.900)
	$\log(P_{middle,jt})$	-9.643**	-4.702**	-6.929***	-8.795***	-7.574***	-7.094***
		(4.348)	(1.885)	(1.143)	(1.474)	(0.796)	(1.141)
2002	$\log(Y_{middle,jt})$	3.471	1.651	3.866***	4.513***	4.301***	4.032***
		(3.681)	(2.503)	(1.084)	(0.918)	(0.839)	(1.019)
	$\log(P_{middle,jt})$	-4.377	-2.737	-4.971***	-5.715***	-5.531***	-5.138***
		(3.897)	(2.690)	(1.324)	(1.153)	(1.031)	(1.260)
2003	$\log(Y_{middle,jt})$	2.316*	0.258	2.624***	3.698***	3.898***	3.303***
		(1.208)	(2.378)	(0.940)	(0.911)	(0.957)	(0.987)
	$\log(P_{middle,jt})$	-2.918**	-1.144	-3.570***	-4.605***	-4.845***	-4.049***
		(1.451)	(2.559)	(1.188)	(1.108)	(1.107)	(1.185)
2004	$\log(Y_{middle,jt})$	2.648*	0.480	3.479***	3.965***	3.588***	2.733***
		(1.561)	(2.316)	(1.275)	(1.114)	(0.972)	(1.022)
	$\log(P_{middle,jt})$	-3.161*	-1.347	-4.562***	-5.067***	-4.549***	-3.446***
		(1.837)	(2.482)	(1.592)	(1.389)	(1.161)	(1.256)
2005	$\log(Y_{middle,jt})$	2.022*	-1.655	2.454**	2.338***	2.800***	2.278***
		(1.225)	(1.851)	(1.094)	(0.892)	(0.900)	(0.875)
	$\log(P_{middle,jt})$	-2.323*	1.066	-3.204**	-2.979***	-3.504***	-2.939***
		(1.412)	(1.955)	(1.355)	(1.099)	(1.047)	(1.068)
2006	$\log(Y_{middle,jt})$	2.131*	-1.178	2.462**	2.314**	2.711***	2.181**
		(1.225)	(2.234)	(1.148)	(0.955)	(0.995)	(1.024
	$\log(P_{middle,jt})$	-2.302	0.606	-3.253**	-3.009**	-3.363***	-2.791**
		(1.404)	(2.324)	(1.450)	(1.201)	(1.175)	(1.263)
2007	$\log(Y_{middle,jt})$	1.953	-1.304	1.122	1.693	3.214***	2.669**
		(1.286)	(2.704)	(1.527)	(1.044)	(1.178)	(1.098)
	$\log(P_{middle,jt})$	-2.096	0.709	-1.969	-2.392*	-4.075***	-3.490***
		(1.469)	(2.723)	(1.722)	(1.276)	(1.401)	(1.331)

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered (by host country) standard errors in parentheses. Omitting poor class. Low is $y|y \leq \underline{y}$, middle is $y|\overline{y} \leq y > \underline{y}$, high is $y|y > \overline{y}$ for j, t. Controls: host remoteness, host urban concentration, host trade openness, host corporate tax, distance, squared skill difference, percent of population that completed no, primary, secondary, and tertiary schooling, dummies for common language, regional trade agreement, customs union, colonial relationship, parent country, and host region. Sample sizes (year): 661 (1997), 781 (1998), 832 (1999), 895 (2000), 992 (2001), 1013 (2002), 1175 (2003), 1278 (2004), 1264 (2005), 1424 (2006), 1493 (2007)

Bibliography

Acemoglu, D. (2008). *Introduction to Modern Economic Growth.* Princeton University Press, Princeton, NJ.

Acemoglu, D., Gancia, G. A., and Zilibotti, F. (2012). Competing Engines of Growth: Innovation and Standardization. *Journal of Economic Theory*, 147(2):570–601.

Agarwal, J. P. (1980). Determinants of Foreign Direct Investment: A Survey. *Review of World Economics*, 116(4):739–773.

Alesina, A. and Rodrik, D. (1994). Distributive Politics and Economic Growth. *The Quarterly Journal of Economics*, 109(2):465–490.

Ali, S. and Dadush, U. (2012). In Search of the Global Middle Class. The Carnegie Papers.

Antras, P. and Nunn, N. (2009). Firms and the Decision to Invest Abroad: Horizontal FDI. Harvard University Ec 2535.

Balassa, B. (1966). American Direct Investments in the Common Market. *Banca Nazionale del Lavoro Quarterly Review*, 77:121–146.

Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5(1):5–32.

Barro, R. J. and Lee, J.-W. (2010). A New Data Set of Educational Attainment in the World, 1950-2010. Working Paper.

Bernasconi, C. (2013). Income Similarity and Bilateral Trade Flows. Working Paper.

Blonigen, B. A. (2005). A Review of the Empirical Literature on FDI Determinants. *Atlantic Economic Journal*, 33(4):383–403.

Blonigen, B. A. and Piger, J. (2011). Determinants of Foreign Direct Investment. NBER Working Paper 16704.

Bloom, N., Draca, M., and Van Reenen, J. (2011). Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity. NBER Working Paper 16717.

Brainard, S. L. (1997). An Empirical Assessment of the Proximity-Concentration Trade-off Between Multinational Sales and Trade. *The American Economic Review*, 87(4):520–544.

Burenstam-Linder, S. (1961). *An Essay on Trade and Transformation.* Almqvist and Wicksells, Uppsala.

Calderon, C., Chong, A., and Stein, E. (2007). Trade Intensity and Business Cycle Synchronization: Are Developing Countries Any Different? *Journal of International Economics*, 71(1):2–21.

Caves, R. E. (2007). *Multinational Enterprise and Economic Analysis.* Cambridge University Press, New York, NY.

Coe, D. T. and Helpman, E. (1995). International R&D spillovers. *European Economic Review*, 39(5):859–887.

Dalgin, M., Trindade, V., and Mitra, D. (2008). Inequality, Nonhomothetic Preferences, and Trade: A Gravity Approach. *Southern Economic Journal*, 74(3):747–774.

Davidson, W. H. (1980). The Location of Foreign Direct Investment Activity: Country Characteristics and Experience Effects. *Journal of International Business Studies*, 11(2):9–22.

Davis, D. R. (1998). The Home Market, Trade and Industrial Structure. *The American Economic Review*, 88(5):1264–1276.

Dunning, J. H. (1988). *Explaining International Production.* Unwin Hyman, London.

Dunning, J. H. (2000). The Eclectic Paradigm as an Envelope for Economic and Business Theories of MNE Activity. *International Business Review*, 9(2):163–190.

Economist (1998). A Grey Market. print edition.

Egger, P. and Pfaffermayr, M. (2005). Trade, Multinational Sales, and FDI in a Three-factor Model. *Review of International Economics*, 13(4):659–675.

Eichengreen, B. and Irwin, D. A. (1998). *The Regionalization of the World Economy*, chapter The Role of History in Bilateral Trade Flows, pages 33–62. University of Chicago Press, Chicago, IL.

Ekholm, K., Forslid, R., and Markusen, J. R. (2007). Export-platform Foreign Direct Investment. *Journal of the European Economic Association*, 5(4):776–795.

Ernst & Young (2012a). Positioned For Growth. Ernst & Young's 2012 Attractiveness Survey. Russia.

Ernst & Young (2012b). Ready For the Transition. Ernst & Young's 2012 Attractiveness Survey. India.

Fajgelbaum, P. D., Grossman, G. M., and Helpman, E. (2009). Income Distribution, Product Quality, and International Trade. NBER Working Paper 15329.

Fajgelbaum, P. D., Grossman, G. M., and Helpman, E. (2011). A Linder Hypothesis for Foreign Direct Investment. NBER Working Paper 17550.

Falkinger, J. (1990). Innovator-Imitator Trade and the Welfare Effects of Growth. *Journal of the Japanese and International Economies*, 4(2):157–172.

FAO (1995). *The State of Food and Agriculture.* Number 28. FAO Agriculture Series, Rome.

Feenstra, R. C. (1996). U.S. Imports, 1972-1994: Data and Concordances. NBER Working Paper 5515.

Feenstra, R. C. (1997). U.S. Exports, 1972-1994, with State Exports and other U.S. Data. NBER Working Paper 5990.

Feenstra, R. C. (2004). *Advanced International Trade.* Princeton University Press, Princeton, NJ.

Feenstra, R. C. and Rose, A. K. (2000). Putting Things in Order: Trade Dynamics and Product Cycles. *The Review of Economics and Statistics*, 82(3):369–382.

Fieler, A. C. (2011). Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation. *Econometrica*, 79(4):1069–1101.

Flam, H. and Helpman, E. (1987). Vertical Product Differentiation and North-South Trade. *The American Economic Review*, 77(5):810–822.

Foellmi, R., Hepenstrick, C., and Zweimüller, J. (2011). Non-homothetic preferences, parallel imports and the extensive margin of trade. CEPR Discussion Papers 7939.

Foellmi, R. and Zweimüller, J. (2006). Income Distribution and Demand-Induced Innovations. *The Review of Economic Studies*, 73:941–960.

Forbes Magazine (2012). The Rise of the Chinese Consumer.

Galor, O. and Zeira, J. (1993). Income Distribution and Macroeconomics. *The Review of Economic Studies*, 60(1):35–52.

Gibson, P., Whitley, D., and Bohman, M. (2001). Profiles of Tariffs in Global Agricultural Markets. U.S. Deparment of Agriculture. Agricultural Economic Report No. 796.

Godley, A. (2001). The Global Diffustion of the Sewing Machine, 1850-1914. *Research in Economic History*, 20:1–45.

Grossman, G. and Helpman, E. (1991a). Endogenous Product Cycles. *The Economic Journal*, 101(408):1214–1229.

Grossman, G. M. and Helpman, E. (1991b). *Innovation and Growth in the Global Economy.* MIT Press, Cambridge, MA.

Grossman, G. M. and Lai, E. L.-C. (2008). Parallel Imports and Price Controls. *The RAND Journal of Economics*, 39(2):378–402.

Hallak, J. C. (2010). A Product-Quality View of the Linder Hypothesis. *The Review of Economics and Statistics*, 92(3):453–466.

Helpman, E. (2006). Trade, FDI, and the Organization of Firms. *Journal of Economic Literature*, 44(3):589–630.

Helpman, E., Melitz, M. J., and Rubinstein, Y. (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. *The Quarterly Journal of Economics*, 123(2):441–487.

Helpman, E., Melitz, M. J., and Yeaple, S. R. (2004). Exports versus FDI with Heterogeneous Firms. *The American Economic Review*, 94(1):300–316.

Heston, A., Summers, R., and Aten, B. (2012). Penn World Table Version 7.1. Center for International Comparisons for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hirsch, S. (1967). *Location of Industry and International Competitiveness.* Oxford University Press, London.

Hirsch, S. (1975). The Product Cycle Model of International Trade - A Multi-Country Cross-Section Analysis. *Oxford Bulletin of Economics and Statistics*, 37(4):305–317.

Horstmann, I. J. and Markusen, J. R. (1996). Exploring New Markets: Direct Investment, Contractual Relations and the Multinational Enterprise. *International Economic Review*, 37(1):1–19.

Hunter, L. (1991). The Contribution of Nonhomothetic Preferences to Trade. *Journal of International Economics*, 30(3-4):345–358.

KPMG (2003). The Grey Market.

KPMG (2008). Effective Channel Management is Critical in Combating the Gray Market and Increasing Technology Companies' Bottom Line: KPMG Gray Market Study Update.

Krugman, P. (1979). A Model of Innovation, Technology Transfer, and the World Distribution of Income. *The Journal of Political Economy*, 87(2):253–266.

Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *The American Economic Review*, 70(5):950–959.

Kugler, M. and Zweimüller, J. (2005). National Income Distributions and International Trade Flows. Working Paper.

Leontief, W. (1953). Domestic Production and Foreign Trade: The Capital Position Re-examined. *Proceedings of the American Philosophical Society*, 97(4):332–349.

Leontief, W. (1956). Factor Proportions and the Structure of American Trade: Further Theoretical and Empirical Analysis. *The Review of Economics and Statistics*, 38(4):386–407.

Limao, N. and Venables, A. J. (2001). Infrastructure, Geographical Disadvantage, Trnasport Costs, and Trade. *The World Bank Economic Review*, 15(3):451–479.

Mahlstein, K. (2010). Die Rolle der ausländischen Direktinvestitionen im Agrarsektor. *Die Volkswirtschaft*.

Markusen, J. R. (1986). Explaining the Volume of Trade: An Eclectic Approach. *The American Economic Review*, 76(5):1002–1011.

Markusen, J. R. (2002). *Multinational Firms and the Theory of International Trade*. MIT Press, Cambridge, MA.

Markusen, J. R. (2010). Putting Per-Capita Income Back into Trade Theory. NBER Working Paper 15903.

Markusen, J. R. and Maskus, K. E. (2002). Discriminating Among Alternative Theories of the Multinational Enterprise. *Review of International Economics*, 10(4):694–707.

Markusen, J. R. and Venables, A. J. (2000). The Theory of Endowment, Intra-industry and Multi-national Trade. *Journal of International Economics*, 52(2):209–234.

Martinez-Zarzoso, I. and Vollmer, S. (2011). Bilateral Trade Flows and Income-Distribution Similarity. Working Papers on International Economics and Finance.

Maskus, K. E. and Chen, Y. (2004). Vertical Price Control and Parallel Imports: Theory and Evidence. *Review of International Economics*, 12(4):551–570.

Matsuyama, K. (2000). A Ricardian Model with a Continuum of Goods under Nonhomothetic Preferences: Demand Complementarities, Income Distribution, and North-South Trade. *The Journal of Political Economy*, 108(6):1093–1120.

Matsuyama, K. (2002). The Rise of Mass Consumption Societies. *The Journal of Political Economy*, 110(5):1035–1070.

Mayer, T. and Zignago, S. (2001). Notes on CEPII's distance measures: The GeoDist database. CEPII Working Paper 25.

McKinsey Global Institute (2010). Lions on the Move: The Progress and Potential of African Economies.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.

Milanovic, B. and Yitzhaki, S. (2002). Decomposing World Income Distribution: Does the World Have a Middle Class? *Review of Income and Wealth*, 48(2):155–178.

Mitra, D. and Trindade, V. (2005). Inequality and Trade. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 38(4):1253–1271.

Muhammad, A., Seale, J. L. J., Meade, B., and Regmi, A. (2011). International Evidence on Food Consumption Patterns: An Update Using 2005 International Comparison Program Data. TB-1929. U.S. Department of Agriculture, Econ. Res. Serv.

Mullor-Sebastian, A. (1983). The Product Life Cycle Theory: Empirical Evidence. *Journal of International Business Studies*, 14(3):95–105.

Murphy, K. M., Shleifer, A., and Vishny, R. (1989). Income Distribution, Market Size, and Industrialization. *The Quarterly Journal of Economics*, 104(3):537–564.

OECD (1999). OECD Benchmark Definition of Foreign Direct Investment. Third Edition. OECD Publishing.

OECD (2012). OECD International Direct Investment Statistics. OECD Publishing.

OECD (2013). FDI in Figures. OECD Publishing.

Osepchuk, J. M. (1984). A History of Microwave Heating Applications. *IEEE Transactions on Microwave Theory and Techniques*, 32(9):1200–1224.

Perotti, R. (1996). Growth, Income Distribution, and Democracy: What the Data Say. *Journal of Economic Growth*, 1(2):149–187.

Persson, T. and Tabellini, G. (1994). Is Inequality Harmful for Growth? *The American Economic Review*, 84(3):600–621.

Posner, M. (1961). International Trade and Technical Change. *Oxford Economic Papers, New Series*, 13(3):323–341.

Ramondo, N., Rappoport, V., and Ruhl, K. J. (2011). Horizontal vs. Vertical FDI: Revisiting Evidence from U.S. Multinationals. Working Paper.

Ravallion, M. (2010). The Developing World's Bulging (but Vulnerable) Middle Class. *World Development*, 38(4):445–454.

Ray, D. (1998). *Development Economics*. Princeton University Press, Princeton, NJ.

Rodriguez, F. and Rodrik, D. (2001). *NBER Macroeconomics Annual 2000*, volume 15, chapter Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence, pages 261–338. MIT Press, Cambridge, MA.

Romer, P. M. (1990). Endogenous Technological Change. *The Journal of Political Economy*, 98(5):71–102.

Santos Silva, J. and Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Scaperlanda, A. E. and Mauer, L. J. (1969). The Determinants of U.S. Direct Investment in the E.E.C. *The American Economic Review*, 59(4):558–568.

Simonovska, I. (2010). Income Differences and Prices of Tradables. NBER Working Paper 16233.

Stobaugh, R. B. J. (1969). Where In the World Should We Put that Plant? *Harvard Business Review*, 47(1):129–136.

Stokey, N. L. (1991). The Volume and Composition of Trade Between Rich and Poor Countries. *The Review of Economic Studies*, 58(1):63–80.

Tellis, G. J., Stremersch, S., and Yin, E. (2003). The International Takeoff of New Products: The Role of Economics, Culture, and Country Innovativeness. *Marketing Science*, 22(2):188–208.

Thurow, L. C. (1987). A Surge in Inequality. *Scientific American*, 256(5):30–37.

Tirole, J. (1988). *The Theory of Industrial Organization.* MIT Press, Cambridge, MA.

UCLA: Statistical Consulting Group (2013). Introduction to SAS. `http://www.ats.ucla.edu/stat/sas/`.

UNCTAD (2012). World Investment Report 2012.

United Nations Statistics Division (2012). Industrial Commodity Production Statistics Database. `http://unstats.un.org/unsd/industry/ics_intro.asp`.

United States Census Bureau (2013). Poverty Thresholds. `https://www.census.gov/hhes/www/poverty/data/threshld/index.html`.

UNU-WIDER (2008). World Income Inequality Database. Version 2.0c.

Vernon, R. (1966). International Investment and International Trade in the Product Cycle. *The Quarterly Journal of Economics*, 80(2):190–207.

Vernon, R. (1979). The Product Cycle Hypothesis in a New International Environment. *Oxford Bulletin of Economics and Statistics*, 41(4):255–267.

Wall Street Journal (2013). Indonesia Foreign Direct Investment Hits. The Wall Street Journal.

Wells, L. T. J. (1966). *Product Innovation and Directions of International Trade.* Harvard University Press, Cambridge, MA.

Wells, L. T. J. (1969). Test of a Product Cycle Model of International Trade: U.S. Exports of Consumer Durables. *The Quarterly Journal of Economics*, 83(1):152–162.

Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach.* Cengage Learning EMEA, 4 edition.

World Bank (2007). Global Economic Prospects: Managing the Next Wave of Globalization. Oxford University Press for the World Bank.

World Bank (2012). World Development Indicators.

Curriculum Vitae

2008 - 2013	Research associate and doctoral student Chair of Macroeconomics and Labor Markets (Prof. Dr. Josef Zweimüller) Department of Economics, University of Zurich
2006 - 2008	Master of Arts UZH (Lizentiat in Ökonomie) University of Zurich
2002 - 2005	Diploma FHNW (Diplom Betriebsökonomie) University of Applied Sciences and Arts Northwestern Switzerland FHNW
1981	Born on Sunday, January 18, in Brugg